Session A1: SCM 1

·Day1: Sep. 15 (Wed.)

·Time: 09:00 - 10:20

·Chair: Milorad Vidović

·Room: Camellia, 5F

The 1st International Conference on Logistics and Maritime Systems

September 15th–17th, 2010 Busan, Korea

COORDINATION OF A PUSH-PULL PRINCIPLE LOGISTICS NETWORK BY OPTIMIZING MATERIAL-PULL; APPLYING GENETIC ALGORITHM

Afshin Mehrsai*, Michael Teucke, Bernd Scholz-Reiter

Department of Production Engineering
University of Bremen-BIBA
c/o BIBA GmbH, Hochschulring 20, 28359
Bremen, Germany
e-mail: (meh|tck|bsr)@biba.uni-Bremen.de

Abstract: Strategies for material flow in logistic networks have a crucial effect on the whole operations. Generally, two types of material flow have been introduced; material push and material pull. Flow of material could be totally push, totally pull, or a combination of both strategies. For hybrid push-pull and totally material pull some specific optimization factors can be considered for higher overall performance. A logistic network scenario, with a push-pull system, is set up to show this impression. Numbers of pallets circulating within the system and the carried lot-sizes are considered to get optimized, considering other dynamic variables. This facilitates a smooth flow between the push and pull sides. Conwip system pulls materials regarding current demands. For dealing with existing complexities, genetic algorithm, as a heuristic search method for solving complex problems, is employed. The better performance is experimented by setting up a discrete- event simulation model and doing several tests.

1. INTRODUCTION

Nowadays, after shifting from simple companies towards supply chains and correlated supply networks, they require more complex logistics processes (Schonsleben, 2000). Under the pressure of global competition, changing business environment, mass customized products, and transient demands, not only the individual plants, but rather logistics networks have acquired decisive rolls for excellence. In fact, planning and control of material flows within single factories as well as supply networks is one of the most complex tasks of current logistics. The complexities accompanied with collaborative logistics networks are the consequence of the paradox in integrating members, while they have their own requirements and performances (Holmstrom and Framling, *et al*., 2002). Consistently, material flows inside shop-floors beside integration and coordination of flows between logistics' members have engaged the most planning and controlling potentials in industries.

Today, concerning the changing business environment, more flexible systems are required to respond to customers' demands, more quickly (Suri, 1998). To be responsive to customers is an inevitable factor for sustaining in such markets (You, and Grossmann, 2008). Consequently, more agile systems are configured to cover the needs of businesses, like flexible manufacturing systems, hybrid systems, distributed and autonomous control systems (Scholz-Reiter, and Freitag, 2007), (Zolfaghari and Roa, 2006), (Yalcin and Namballa, 2005). These all are particularly necessary for moving toward mass customization strategy, despite initial successes of mass production. Specially, mass customization is the facing or prospective situation of many businesses. Although, the mass production system follows push material concept by more production higher benefit, for some other approaches pull of material is more consistent (Ozawa, 2006). Generally saying, undergoing one of the both push or pull strategy has a direct effect on the performances of overall logistics processes.

However, to some extent, businesses are confronted with continuous changing conditions, called dynamics, which are supposed to be handled by more intelligent strategies. For instance, mass customized products enforce supply networks to follow the make-to-order (MTO) or engineer-to-order (ETO) production/logistics strategies to comply with individual demands. These recent production strategies burden more pressure on networks to operate based on real demand and at the right time. The real time operation reflects an agile pull principle strategy. Nevertheless, push strategy e.g., MRP, results better when high variety exists and demand fluctuates, but still is predictable (Rocky and Sridharan, 1995). In contrary, pull strategy e.g., Kanban, complies better with rather stable demand and low variety in products (Scholz-Reiter, Mehrsai and Görges, 2009). In addition, different material flow strategies have their own benefits and drawbacks. For example, shifting from totally push system to Kanban system, with fully pull concept, may have some shortcomings in facing uncertainties. Therefore, a clever solution for dealing with such conditions is to employ advantages of several material flow control strategies (Scholz-Reiter and Mehrsai, 2009a). For this purpose, some mutual systems have been suggested, e.g., Conwip, Polca, G-Polca, and in a broader scale Leagility, to overcome those problems (Spearman and Woodruff, 1990), (Fernandes and do Carmo-Silva, 2006). These systems compensate the potential weaknesses of sole strategies and hinder them from getting failed or overproduction. Furthermore, employing this idea supports the needs in terms of responsiveness, quickness, flexibility, reliable delivery, agility as well as leanness in logistics (Scholz-Reiter and Mehrsai, 2009a).

In general, most of the above stated systems have been emphasized on production and shop-floor logistics while they could be effectively applied by supply chains and logistics networks. For instance, within a logistics network with application of push material up to a specific point, called decoupling point (Sun *et al*., 2008), material flow could follow push planning and control systems, like MRP. Whereas downstream of the network can have a pull or hybrid push-pull system. This is specifically useful for benefiting from both control concepts. Nevertheless, coordination of downstream pull system with the upstream push one is a challenging issue.

This paper deals with the importance of optimized number of conveying means (pallet) and their lot-size in the pull environment. The conveying means represent items for controlling Conwip (constant work in process) system. Here, a discrete-event simulation scenario of a hybrid logistics network is developed to emphasize the roll of coordinated hybrid push-pull network, facing dynamics in its own processes. Considered dynamics are stochastic demands, fluctuating supply, and uncertain processing time of operations that resemble real-world problems. This hybrid concept contributes to improving logistics performance measures, e.g., throughput time (TPT), throughput (TP), responsiveness, utilization, and work in process (WIP), (Gunasekaran, Patel and McGaughey, 2004). However, several stochastic variables affect materials control in shop-floors (Petroni and Rizzi, 2002) and, to some extent, throughout logistics networks. Hence, the optimization problem has a complex and non-linear behavior. For this purpose, genetic algorithm (GA) as a stochastic optimization method is chosen for solving this problem (Wang and He, 2009). Additionally, by exploiting advantages of fuzzy set theory, the improvement in recognizing uncertain processes is illustrated.

The main goal of this study is to show the privileges of improving logistics by approximating optimum values of just few key factors. The target is to show suitability of the network's material flow control strategy, GA optimization process, and superiority of fuzzy sets application in better distinguishing uncertain process times. The GA and Fuzzy methods are applied for realizing this goal. The rest of the paper is organized as follows: next section proceeds with briefly explaining the types of material flow control systems. Section 3 refers to genetic algorithm and its approach in solving stochastic optimization problems. Section 4 shortly introduces fuzzy set theory application in production and logistics. The logistics network scenario is given in section 5, and the problem statement is displayed in section 6. Experimental results are depicted on section 7. The summary and further suggestion is written in section 8.

2. MATERIAL FLOW CONTROL SYSTEM

Generally, material flow control systems can be classified as push and pull mechanisms (Fernandes and do Carmo-Silva, 2006). These two principles trigger a spectrum of strategies to produce and flow materials between two locations, may be two workstations or, to some extent, two members of a supply network.

2.1 Material push control

Initially, material flow systems used to control flows by push principle. It means, pushing material to next processing steps as soon as its process is finished at the current step. In this case, if the line or workstations are not balanced together, WIP are collected everywhere and overproduction is the consequence of this system. However, this control system is more suitable for mass production and make-to-stock (MTS) strategies with balanced lines. Nevertheless, line balancing is a challenging issue by itself when the system is instable. Material requirement planning (MRP) is a well-known method categorized into push control system. However, some authors partially classified drum-buffer-rope (DBR), starvation avoidance (SA), G-Polca, and even Conwip, as push control systems (Fernandes and do Carmo-Silva, 2006), (Pahl, Voss, and Woodruff, 2005), (Germs and Riezebos, 2009).

2.2 Material pull control

Material pull control has been originated by Toyota production system (TPS) as Kanban (Chan, 2001). This concept made a breakthrough in Toyota and later other imitating industries for a long time. Nevertheless, this control principle encounters some difficulties when demand is oscillating and uncertainty is inherently by processes. Polca (Suri and Krishnamurthy 2003) and partly Synchro-MRP (Geraghty and Heavey, 2004) are categorized in pull control system as well. However the tow mechanisms exploit some aspects of push. Therefore, some hybrid control systems have been considered to resolve the dilemma accompanied with pure pull systems. Among them, Conwip, and Generic-Polca could be mentioned.

2.3 Hybrid push-pull control

Although MRP is a central control system, which pushes materials to downstream of consumption point, but, in contrary, pull principles are categorized in distributed control systems. Pull systems work based on WIP limitation and current demand of the local working area. Concurrent application of the both push-pull systems gives a twofold view to a seamless controlling of material flows and streamlines the flows. Firstly, employment of push (by central control) defines the release dates of processing in global context. Meanwhile, pull control operates based on local situations of WIP. By doing so, global and local factors interact with each other. This approach enhances coordination of the entire logistics system facing dynamics.

However, simultaneously employment of push and pull systems is not necessarily required. Inspired by shop-floors control, some suggested control strategies for logistics networks can be sorted as follows:

- Dividing the entire network into two parts as push and pull, which is broadly discussed as Leagile supply chains (Mason-Jones, Naylor and Towill, 2000).
- Employment of both push and pull control systems simultaneously within each member of the network or throughout the whole network, like: Conwip and G-Polca (this type needs high flexibility entirely which is subject to have distributed and intelligent control system).
- Inspired by Polca, dividing the network into paired-cells and applying the material release date by push system as well as WIP limitation by pull cards.

In these listed options, dispatching rules and workload balancing are still challenging. However, in this paper just the first proposed option is analyzed. Downstream with pull material flow is optimized to coordinate the collision point of push-pull flows. Dispatching rule and balancing workloads are chosen to be done based on an autonomous control system, which is briefly explained in section 5. It follows the bottleneck control rule, but based on autonomous objects' decisions and less queue length (Scholz-Reiter, Görges and Philipp, 2009).

3. GENETIC ALGORITHM

In general, a number of optimization methodologies have been introduced to solve non-linear and NP-hard problems. GA, as a competent evolutionary technique, is a stochastic optimization method based on heuristic procedures (Kubota and Fukuda, 1999). It has been shown that GA is able to find approximately optimum solutions within a fairly quick time. Optimization process of GA starts by randomly generating a population of solutions (individuals), which are in the format of genotype. The specification of a solution can be stored in one or more chromosomes. A chromosome is made of an ordered sequence of single genes. In each gene a single parameter of a coded solution (genotype) is carried. A genotype carries the coded solution, which by decoding that to the original solution is called phenotype. The position of a gene in a chromosome is named locus (Schönberger, 2005). Mostly, to codify the solution of a problem, binary-based encoding procedure is selected. Nonetheless, encoding is not limited to binary values, e.g., here, integer values are used.

However, the initial population, which is randomly generated, is subject to get improved to achieve optimum solution. In general, GA has two strong driving engines to produce new solutions without having any knowledge in prior. Selection and adaption operations are done by application of crossover and mutation functions of GA. For crossover function two individuals from a population are considered to get merged and produce either one child (offspring) or two children. There are one-point or multi-point crossover procedures. Moreover, individuals' selection and crossover point on each individual are the factors that influence the quality of the optimum solution and the spent search time. Similarly, mutation is also a function of optimization procedure which avoids local traps. For example, exchanging two genes in an individual or shifting some genes from one locus to other one are two ways of this procedure. Inversion of an individual's genes could be assumed as mutation. The general used framework of GA for this study is as follows:

```
Begin
    t = 0;
    initialize new population P(t);
    evaluate the fitness value of current solutions (individuals)
    while (generation number <7) do
    begin
            t =t+1;
            select individuals for P(t) with higher probability from P(t-1);
            alter (crossover and mutate) individuals in P(t);
```

```
            evaluate P(t);
      end;
end;
```

Furthermore, optimizing elements are based on fitness function values, evolution of individuals, and selection method. Fitness function is an objective function to assess individuals and assign a fitness value to each. According to those fitness values, probabilities of each individual to get selected for the next generation are defined. The best individuals breed the next generation and eliminate the weak performing solutions. In fact, after generating enough new individuals each fitness value of them will be measured. Afterwards, depending on type of selection operator, their selection probabilities, proportional to each other, will be calculated. Weak individuals will be substituted by those with better performing in solving the problem, as parents for next generation. Then, the new population of solutions will be produced by means of reproduction operator. Genetic operators as selection methods can have different mechanisms for selecting parents of the next generation, e.g., roulette-wheel selection, stochastic remainder selection, stochastic universal sampling, and tournament selection (Cantú-Paz, 2001). For this paper, roulette-wheel selection is used to evaluate the solutions (1). This method has homogeneity in defining probabilities. The roulette-wheel function measures a probability of selection for each individual by getting the mean value of fitness (f) of an individual in proportion to all observations of the fitness values. Equation (1) defines the probability function. Here, N is the number of individuals in current population. Easily, the higher the probability value the more chances have the individuals to be selected.

$$P_i = \frac{f_i}{\sum_{j=1}^{N} f_j} \qquad (1)$$

Although GA has had several achievements in solving complex problems, but its ordinary algorithm does not always lead to an optimum solution, in limited time. Missing and premature convergence and infeasible solutions (individuals) are the threats of the general form of GA. Therefore, some local adaptations and modifications could be integrated to the procedure of producing individuals in order to evade inefficiency and improve the optimum problem search. In this case GA is recognized as Memetic algorithm (MA), (Ong and Keane, 2004). In an optimum seeking search MA tries to evaluate different methods of selection, crossover and mutation, in order to achieve the best solution in a confined time horizon.

4. FUZZY SET

Fuzzy set theory is a powerful tool for characterizing uncertainty and stochastic nature of practical operations in logistics. Practitioners are aware that any human-centered problems or processes e.g., processing times, and due dates, are uncertain (Sakawa and Kubota, 2000), orders comes stochastically, and available information are imprecise. Fuzzy set theory by using fuzzy numbers, their membership functions, and defining fuzzy rules, distinguishes and compromises existing uncertainty as well as imprecision accompanied by processes. It suits to vague or ill-defined problems as well. Specifically, here, uncertain processing times e.g., normal or exponential distribution, in shop-floors or outbound logistics operations causes imprecise decisions over material flow scheduling and control. This problem can be solved by taking into account the fuzzy nature of them and arranging fuzzy rules for better resulting decisions. IF-Then fuzzy rules reflect the policy of the decision maker in terms of the problems' objectives (Petrovic, *et al.*, 2008).

Several shapes can be used for defining membership functions of fuzzy sets that among them are triangular, trapezoidal, Gaussian, and s-curve (Fayad and Petrovic, 2005). Triangular fuzzy membership function, because of its simple arithmetic operations, is usually considered for modeling uncertain processing times. This membership function is represented by a triplet (y^1, y^2, y^3). While y^1 is the lower bound and y^2 is the upper bound of the fuzzy number ($\tilde{y}$) with membership degree of zero ($\mu_{\tilde{y}}=0$), y^2 is the modal point (middle range) with membership degree of one ($\mu_{\tilde{y}}=1$). This type of fuzzy membership is also chosen to be used for the current problem, which is explained later.

5. LOGISTICS NETWORK SCENARIO

A simulation scenario of a logistics network is developed to present the performance improvement in push-pull material flow system, applying GA. Here, it is shown that GA has the ability to coordinate the collision point of push-pull material flows, just by optimizing the pull side material flow. This task is done by using discrete-event simulation approach and offline optimization GA.

Although, optimization by GA has the capability to get done in real-time operations, but this is not the case in this paper. In practice, this can be carried out by application of autonomous learning pallets within a pull principle system.

In pull systems, pallets (or any means of transportations) circulate permanently within logistics systems. Hence, they can be used as pull signals (Scholz-Reiter and Mehrsai, 2009b). Pallets have the chance to evaluate the system concurrently and decide what to do for optimizing next steps. Since GA is a global search technique (Ong and Keane, 2004) the optimization process is considered to be offline so that using entire information of the scenario.

The considered simulated network is constructed by three steps of processing plants. In step 1, two plants (P_{11}, P_{12}) are considered to produce three types of raw material in each plant. Each type of raw material has to get assembled with its counterpart from other plant, in the next step. The step 2 has two assembly plants (P_{21}, P_{22}) which have the same processing capability. Therefore, the plants in step 1 are fully connected to the plants in step 2 and products will be allocated to them based on queue length and bottleneck control concept. The plants in step 2 will transfer their assembled products (which are now just three types) to the final plant called OEM. From sources up to the entrance of OEM products are push and just inside OEM pull principle is used. Figure 1 shows the view of the network. For each of the three types of products, based on the exponential distribution time intervals, one order is triggered to the OEM. Hence, the demand is uncertain based on time intervals between each order for each product. Equation (2) represents the used negative exponential distribution.

$$f(x) = \frac{1}{\beta} \cdot \exp\left(-\frac{x}{\beta}\right) \qquad (2)$$

Here, the mean value ($\beta=1/\mu$) is assumed a bit bigger than the mean supply rate (i.e., $\beta \approx$ 2:40 min) to collect WIP in entrance inventory. The variance value of (2) is equal to β^2.

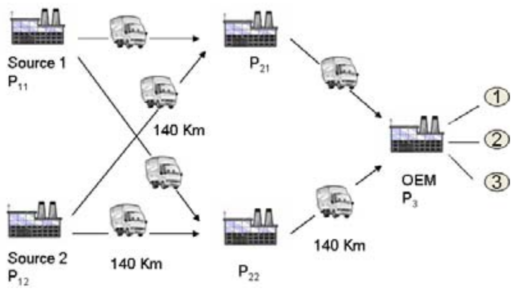


Figure 1. Exemplary push-pull network, with lasting each round trip 4 hours for transporters

It is noticeable that inside each plant, except the assembly ones in step 2, a 3x3 matrix of workstations is devised. The matrix configures three similar production lines in parallel, which are fully coupled to each other after each workstation. The purpose of this fully coupled system is to simulate a high flexible logistics system with capability of applying autonomous logistics objects. Actually, autonomous objects by collecting local information about successive queues (buffers in front of each station) and using bottleneck control rule, decide which route has the least waiting time. This control system has been discussed in previous papers (Scholz-Reiter and Freitag, 2007), (Scholz-Reiter, Mehrsai and Görges, 2009), (Scholz-Reiter and Mehrsai, 2009a), (Scholz-Reiter, Görges and Philipp, 2009). Table 1 shows the processing times of each workstation inside each plant. For OEM the processing times are considered stochastic with normal distribution. Standard deviation for the distribution equals to ($\sigma = \mu/10$). The mean of processing times is equal to the mean intervals of emerging products. The simulation runs for 80 days each 24 hours.

Table 1. Processing times for each product on each line

	Processing times [h:min] for each plant						
Plant	**P_{11}; P_{12}**			**P_{21}; P_{22}**	**P_3 (OEM)**		
	Line						
	Deterministic value				**Mean value (μ)**		
Product	**1**	**2**	**3 1 1**			**2**	**3**
Type 1	2:00	3:00	2:30	0:50	2:00	2:40	2:20
Type 2	2:30	2:00	3:00	0:50	2:20	2:00	2:40
Type 3	3:00	2:30	2:00	0:50	2:40	2:20	2:00

The used flow strategy is as follows: in step 1, materials are discharged to the network based on release dates according to normal distribution. The normal distribution is arbitrarily assumed for time intervals between each release with $\mu=50$ min and $\sigma=5$ min. Then the three product types are randomly released to the system. Shop-floors control at upstream with push material are assumed to get balanced by autonomous products. This autonomy avoids congestions in shifting bottlenecks as the consequence of uncertainty in release dates. In contrast, at downstream, demand orders are enquired in stochastic manner with exponential time intervals. The downstream control of pull material performs with a Conwip-like system. Availability of circulating pallets at the entrance of OEM is the signal of stated demand. Pallets do the duty of Conwip cards. They are pushed by autonomous controlled to the downstream of the shop-floor after picking up their respective products. Eventually, uncertainties in replenishment push and order pull result in a chaotic performance in the collision (decoupling) point of push-pull that its coordination is supposed to get optimized.

6. PROBLEM STATEMENT

6.1 General Problem

For explaining the optimization problem, on the one side, flexibility issues beside rate of material push are considered. On the other side, upcoming orders in pull section should be seen. The supposed flexibilities for the experiments are autonomous control for pallets in routing as well as flexible lot-sizes and number of cycling pallets. Here, a great complexity is the arrangement of empty available pallets for upcoming demand on time. Since the flow of pushed materials and upcoming demands are uncertain and unstable, number of Conwip cards (pallets here), and lot-sizes, are the optimizing factors for the material flow problem. However, their exact contribution to the objective is mathematically unknown in advance. This is a strong reason for employing simulation and GA to solve this problem. Here, the multi-objective problem is to minimize average local through put time (ATPT), Average global TPT (AGTPT), and the OEM entrance's inventory (WIP), as well as to maximize total deliveries (TD) to final customer. The used notations are as follows: Product type is denoted by P; p=1…P, P=3.Time is denoted by t; t=0…T, T=80 Days. $ALTPT_p$ denotes average local throughput time (in OEM at T). $AGTPT_p$ denotes average global throughput time at T. Total delivery of product p at time T is denoted by TD_p. Maximum OEM inventory is denoted by WIP_p. Importance weights are denoted by α, δ; may be chosen arbitrarily by decision maker.

Since the objectives of the problem are min and max as well as two different units (time and number), a suitable solution for making the objectives homogeneous is to transform them into corresponding satisfaction degrees. A good application of this solution is recently presented by (Petrovic *et al.*, 2008). All objectives can be simply mapped into a linear membership function, as satisfaction degree within the range of [0 1], just by defining the upper and lower bound of the objective by decision maker. Then by using fuzzy rules, minimum operator, or simply using the average aggregation operator for the satisfaction degrees, one unique satisfaction degree as entire objective, will be achieved. However, explaining this transformation process is not the goal of this paper. Instead here, all single objective values are converted into real numbers without any units.

Here, the target is to show the suitability of the network's material flow control strategy, GA optimization process, and superiority of fuzzy sets application in better distinguishing uncertain process times. For this purpose by using GA, it just needs to minimize a fitness function. This fitness function can be easily formulated as (3) that is considered as a unique unitless function.

$$\text{Min}\sum_{P}\left(\frac{\frac{ALTPT_p}{P}+\frac{AGTPT_p}{P}+WIP_p\times\alpha}{\frac{TD_p}{P}\times\delta}\right) \tag{3}$$

6.2 Solution Characteristic

Application of fuzzy sets is just used in control level inside autonomous pallets. The triangular fuzzy numbers for approximating the processing time of stations are considered as (1:48, 2:00, 2:12), (2:06, 2:20, 2:34), (2:24, 2:40, 2:56) for three product types. These values are exerted to recognize uncertain waiting times and upon them choosing the best route with the least waiting time. In order to compare fuzzy waiting times of parallel stations the pallets have to use some ranking criteria for comparing the fuzzy waiting values (4), (Sakawa and Kubota, 2000).

$$C_1(\tilde{Y})=\frac{y^1+2y^2+y^3}{4},C_2(\tilde{Y})=y^2,C_2(\tilde{Y})=y^3-y^1 \tag{4}$$

The specific GA for this problem is supposed to optimize the fitness value, which is a minimization function. Fitness values (called observation) for each individual, of the first generation, are first calculated by (3). Then based on (1) the selection's probabilities for each individual are found. Derived from the probability values, randomly, ten individuals will be chosen for crossover and mutation. The first created generation is configured by 10 individuals. Crossover function has probability of 0.8 and mutation of 0.1. In the next generations the procedure is the same. GA by using (1) defines selection probabilities and then takes each couple of individuals to breed two children by means of crossover and then mutation. From the second generation to the final one, the populations are combined of 20 individuals. In the second generation, again 10 individuals with higher probabilities are selected to become parents for generating new children. This repetitive procedure runs up to the termination value which is set as the number of generations (6 here). All individuals in new generation are evaluated unless they have been seen in the previous generations. Totally, GA experimented 110 ($10+5\times20=110$) individuals consist of the couple of lot-size and pallets number values, for seeking the excellence.

7. EXPERIMENTAL RESULTS

In this section, two alternative outcomes of the simulation are displayed by surface graphs. The first graph (Figure 2a) depicts the GA search for three dimensions as lot-size, number of pallets, and fitness value, when the autonomous control considers no fuzzy sets. In this case the processing time are assumed to be crisp by using the mean value of distributions. Obviously, the search procedure is uneven and may be imprecise.

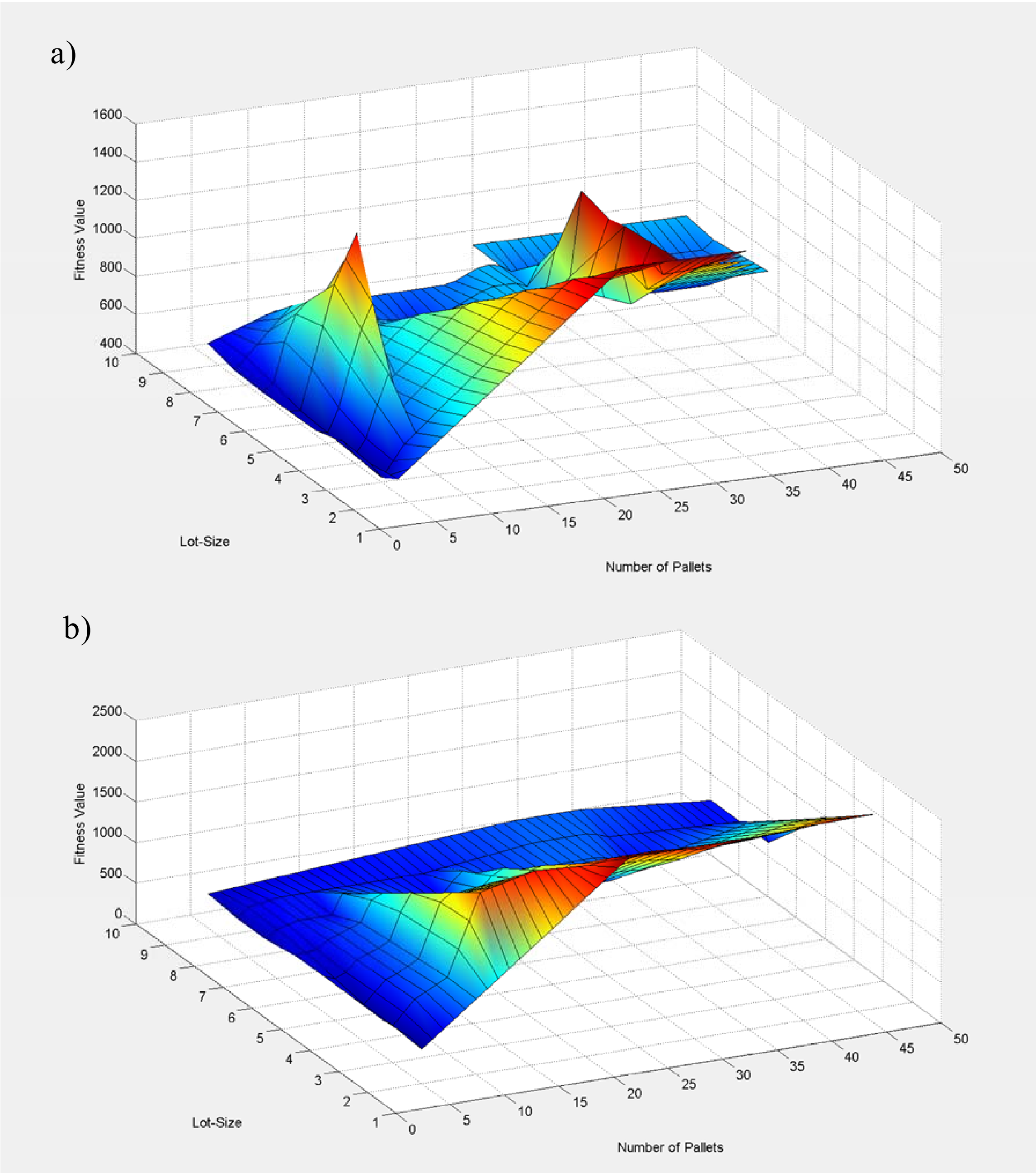


Figure 2. a) GA search without using fuzzy set theory, b) GA search with using fuzzy set theory

Here, the simulation runtime was 6:42 minutes and the best combination is recognized as 5 pallets and 3 lot-size with the best fitness value of 422.82 units.

On the other hand, the Figure 2b shows the same outputs when the system considers fuzzy nature of processes and takes into account the ranking criteria. In this case with fuzzy sets, runtime was 7:34 minutes and the best fitness value was 415.47 units. The Best combination is found as 5 pallets and 2 lot-size. Obviously, the fitness value is dropt here by about 7.5 units and the lot-size decreased by one. As it can be seen, the graph in Figure 3 has a homogeneous shape, whereas the one in Figure 2 has a chaotic shape. This proves that when the system is uncertain using fussy numbers and operations performs better than crisp values with their conventional calculations. Finally, application of GA gave the opportunity to decision maker for adapting its system to optimality through a broad range of available searched values.

8. SUMMARY

To sum up the work, in the first steps some introductions about material flow control strategies were given. The existing strategies complying with material pull and material push were briefly explained and some examples were mentioned. The most emphasis of the paper was on the advantages of applying a hybrid control system out of push and pull concepts. This was inspired by Polca, Conwip, and other comparable systems to control a smooth and robust flow of materials in an instable system, regarding uncertainty and fluctuating demand. Afterwards, for optimizing push and pull flows, GA as a global optimization method was concisely described and its procedure to get the optimum solution have been defined. Following that an abstract application of fuzzy set theory for defining uncertain processes was given. It was shown, that GA has the capability to coordinate the both side of a stochastic material flow in collision point of material push and material pull. Finally, in the last section the results of the experiments were depicted and shortly proved the improvement in the system performance, by optimizing the pull system according to the pushed material flow. It was displayed by graphs that the operations with fuzzy values perform better. However, the optimization was not run in real-time of material flow, but by assistance of simulation with several offline experiments. This real-time optimization is the subject of the following studies. Through applying some learning methodologies e.g., artificial neural network (ANN), Lamarckian learning for improving GA (Ong and Keane, 2004), employed by own autonomous objects, pallets may get the capability to learn and improve the performance of the local and global system in real-time. It can be realized just by collecting real-time experiences. Thereby even these logistic objects can switch between push and pull performance in real-time based on current situation. This research project is entitled as “learning pallets” by the authors.

ACKNOWLEDGEMENTS

The current study is supported by the International Graduate School for Dynamics in logistics at Bremen University. The authors wish to thank the anonymous colleagues for their reviews.

5. REFERENCES

Cantú-Paz, E. (2001). Migration Policies, Selection Pressure, and Parallel Evolutionary Algorithms. *Journal of heuristics, 7/4*:311-334.

Chan, F. (2001). Effect of Kanban Size on Just-In-Time Manufacturing Systems. *Journal of Materials Processing Tech., 116/2-3*:146-160.

Fernandes, N. and do Carmo-Silva, S. (2006). Generic POLCA—A Production and Materials Flow Control Mechanism for Quick Response Manufacturing. *International Journal of Production Economics, 104/1*:74-84.

Fayad, C. and Petrovic, S. (2005). A Fuzzy Genetic Algorithm for Real-World Job-Shop Scheduling. *M. Ali and F. Esposito (Eds.), Innovations in Applied Artificial Intelligence, Lecture Notes in Artificial Intelligence 3533.* Springer-Verlag Berlin Heidelberg, Germany, 524-533.

Gunasekaran, A., Patel, C. and McGaughey, R. (2004). A Framework for Supply Chain Performance Measurement. *International Journal of Production Economics, 87/3*:333-347.

Germs, R., Riezebos, J. (2009). Workload Balancing Capability of Pull Systems in MTO Production. *International Journal of Production Research, 48:* 8, 2345- 2360.

Geraghty, J. and Heavey, C. (2004). A Comparison of Hybrid Push/Pull and CONWIP/Pull Production Inventory Control Policies. *International Journal of Production Economics, 91/1*:75-90.

Holmstrom, J, Framling, K., Tuomi, J., Karkkainen, M. and Ala-Risku, T. (2002). Implementing Collaboration Process Networks. *International Journal of Logistics Management, 13/2*:39-50.

Kubota, N. and Fukuda, T. (1999). Structured Intelligence for Self-Organizing Manufacturing Systems. *Journal of Intelligent Manufacturing, 10/2*:121-133.

Mason-Jones, R., Naylor, B. and Towill, D. (2000). Engineering the Leagile Supply Chain. *International Journal of Agile Management Systems, 2/1*:54-61.

Ozawa, T. (2006). *Institutions, Industrial Upgrading, and Economic Performance in Japan: the'flying-geese'paradigm of catch-up growth*, Edward Elgar Pub, Northampton.

Ong, Y. and Keane, A. (2004). Meta-Lamarckian Learning in Memetic Algorithms, *IEEE Transactions on Evolutionary Computation, 8/2*:99-110.

Petrovic, S., Fayad, C., Petrovic, D., Burke, E. and Kendall, G. (2008). Fuzzy Job Shop Scheduling with Lot-sizing. *Annual Operation Research, 159*: 275-292.

Pahl, J., Voss, S. and Woodruff, D. (2005). Production Planning with Load Dependent Lead Times. *4OR: A Quarterly Journal of Operations Research, 3/4*:257-302.

Petroni, A. and Rizzi, A., (2002). A fuzzy logic based methodology to rank shop floor dispatching rules. *International Journal of Production Economics, 76/1*:99-108.

Rocky, W. and Sridharan, V. (1995).Linking Manufacturing Planning and Control to the Manufacturing Environment. *Integrated Manufacturing Systems, 6/4*: 36-42.

Schonsleben, P. (2000). With Agility and Adequate Partnership Strategies Towards Effective Logistics Networks. *Computers in Industry, 42/1*:33-42.

Suri, R. (1998). *Quick response Manufacturing: a Companywide Approach to Reducing Lead Times*, Productivity Pr.

Scholz-Reiter, B. and Freitag, M. (2007). Autonomous Processes in Assembly Systems. *CIRP Annals-Manufacturing Technology, 56/2*: 712-729.

Scholz-Reiter, B., Mehrsai, A. and Görges, M. (2009). Handling Dynamics in Logistics - Adoption of Dynamic Behaviour and Reduction of Dynamic Effects. *AIJST-Asian International Journal of Science andTechnology Production and Manufacturing Engineering (AIJSTPME), 2/3*: 99-110.

Scholz-Reiter, B. and Mehrsai, A. (2009a). Integration of Lean-Agile Experiments with Autonomy in Supply Chains. *Proceedings of the 7th International Conference on Manufacturing Research (ICMR09),* University of Warwick, UK, 60-66.

Spearman, M., Woodruff, D. and Hopp, W. (1990). CONWIP: a Pull Alternative to Kanban. *International Journal of Production Research, 28/5*:879-894.

Sun, X., Ji, P., Sun, L. and Wang, Y. (2008). Positioning Multiple Decoupling Points in a Supply Network. *International Journal of Production Economics, 113/2*:943-956.

Suri, R. and Krishnamurthy, A. (2003). *How to Plan and Implement POLCA–A Material Control System for High Variety or Custom-Engineered Products.* Center for Quick Response Manufacturing, Madison: University of Wisconsin,USA.

Scholz-Reiter, B., Görges, M. and Philipp, T.(2009). Autonomously Controlled Production Systems—Influence of Autonomous Control Level on Logistic Performance. *CIRP Annals-Manufacturing Technology, 58/1*:395-398.

Schönberger, J. (2005). *Operational Freight Carrier Planning*. H.-O. Günther and P. v. Beek (Eds.), Advanced planning and scheduling in process industry. Springer, Heidelbelg, Germany.

Scholz-Reiter, B., Mehrsai, A., (2009b). Superior Performance of Leagile Supply Networks by Application of Autonomous Control. *Proceedings of Conference of Advances in Production Management Systems (APMS2009),* Bordeaux, France.

Sakawa, M. and Kubota, R. (2000). Fuzzy Programming for Multiobjective Job Shop Scheduling with Fuzzy Processing Time and Fuzzy Duedate Through Genetic Glgorithms. *European Journal of Operation Research, 120*: 393-407.

Wang, B. and He, S., (2009). Robust Optimization Model and Algorithm for Logistics Center Location and Allocation under Uncertain Environment. *Journal of Transportation Systems Engineering and Information Technology, 9/2*:69-74.

You, F. and Grossmann, I. (2008). Design of Responsive Supply Chains under Demand Uncertainty. *Computers and Chemical Engineering, 32/12*: 3090-3111.

Yalcin, A. and Namballa, R. (2005). An Object-Oriented Simulation Framework for Real-Time Control of Automated Flexible manufacturing systems. *Journal of Computers & Industrial Engineering, 48/1*:111-127.

Zolfaghari, S. and Roa, E. (2006). Cellular Manufacturing versus a Hybrid System: a Comparative Study. *Journal of Manufacturing Technology Management, 17/7*: 942-961.

MIP MODEL AND HEURISTICS FOR INVENTORY ROUTING PROBLEM IN FUEL DELIVERY

Vidović Milorad[1] , Popović Dražen[2]

Faculty of Transport and Traffic Engineering,
University of Belgrade,
Vojvode Stepe 305,
Belgrade, Serbia
e-mails: [1]m.vidovic@sf.bg.ac.rs, [2] d.popovic@sf.bg.ac.rs

Abstract: This paper presents solution approaches to multi period time horizon Inventory Routing Problem (IRP) in fuel delivery, with deterministic consumption at petrol stations. IRP consists of two mutually dependent sub problems, inventory and routing, in Vendor Management Inventory (VMI) environment where suppliers determine quantities and time intervals of deliveries. For solving of IRP we developed mathematical formulation of mix integer programming (MIP) model for purpose of small scale testing. Results were compared with exact heuristic model that was also introduced in this paper. Heuristic approach was developed for purpose of solving large scale problems.

Keywords: IRP, Mix Integer Programming, Heuristics, Fuel Delivery

1. INTRODUCTION

Two of the main components of the supply chain, which have the most significant impact to its performances, are inventories and transportation costs. Although this fact is well known, modeling approaches to supply chain optimization usually consider inventory control and transportation independently, ignoring its mutual impact. However, inter-relationship between the inventory allocation and vehicle routing has recently motivated some authors to model these two activities simultaneously. This practical and challenging logistical problem is known as the integrated Inventory Routing Problem Moin and Salhi (2007). In the IRP idea is to simultaneously solve problems of choosing optimal quantity and time of delivery of goods, as well as the problem of optimal routing of vehicles. The objective is to find balance between inventory and routing costs in such a way to minimize total costs that are incurred by these two segments. The IRP assumes application of Vendor management inventory concept where suppliers determine order quantity and time of its delivery.

Different types and concepts of IRP can be found in available literature. Webb and Larson (1995) observe two types of IRP: tactical (TIRP) and strategic (SIRP). TIRP models considers solving routing problems and their impact on inventory, while SIRP models deals with determination of fleet size (in case when fleet size can not be enlarged in short time period) considering all possible TIRP variants. This paper simultaneously observes both of these IRP types. Concept one-to-many is commonly applied in IRP model where deliveries are made from one depot to many locations as it is case with Betrazzi and Speranza (1997) and Viswanathan and Mathur (1997). Both of these papers have one depot and many delivery locations where deliveries can be realised only in defined frequencies (every two, three, five days, ...). Later paper observes multi-product system. Concept of collecting commodities from many locations to one depot or manufacturing facility (reverse logistics) is of newer date and is explored by Lee et al. (2003). They consider collection of computer parts from different location to one manufacturing facility where stock level per each part cannot drop to zero (due to expensiveness of stopping the production line) and therefore backordering is not allowed.

Another classification of IRP model is by size of planning horizon. Single-period models are usually basis for multi-period models and are used when it is difficult to predict demand in longer period of time at each location. Problem with one-period or short-period models is that they have tendency do transfer as many deliveries as possible to next planning period. If demands can be determined in longer period of time than we can introduce longer planning horizon and obtain better solution.

Demands can be deterministic (which is rare in real cases) and stochastic in which case they can be presented in model by random distributions. Ribeiro and Lourenzo (2003) considers IRP model in one week planning horizon with deterministic and with stochastic demands. Regardless to type and characteristics of IRP, optimal solution for real case study is so far unreachable because of complexity of solution space (main reason lays in routing segment). Therefore many authors propose different heuristic approaches for solving real case IRP models. There are many application areas of IRP concept, from automotive industry, gas companies, food industry, retail of general commodities, etc. For detail insight in IRP see review paper from Moin and Salhi (2007).

This paper is focused on multi period time horizon IRP in fuel delivery, with deterministic consumption at petrol stations. Problem if formulated as mix integer programming (MIP) model. Since routing problem is NP-hard (see

Bramel and Simchi-Levi, 1997), then IRP is also NP-hard because it includes routing problem as one segment of the model and therefore it is impossible to be solved in case of large scale problems. This is the reason why we introduce heuristics. Results of proposed heuristic are compared with solutions obtained by MIP model on the set of test examples of moderate size, which are used then as benchmarks to estimate performances of heuristics solutions.

The paper is organized as follows. Model formulation is given in Section 2. Section 3 contains description of MIP formulation. Description of the proposed heuristic is given in Section 4. Computational results are presented in Section 5. Finally, Section 6 contains concluding remarks and direction for further research.

2. PROBLEM FORMULATION

The IRP problem studied in this paper can be described as multi period deterministic IRP in fuel delivery. Delivery quantities of $\boldsymbol{J}$ fuel types for given set of $\boldsymbol{I}$ petrol stations must be determined through the entire planning horizon. Fuel is transported by one vehicle type that has two compartments. Vehicle can tow trailer with two compartments of the same capacity. Therefore, in one route maximally four compartments ($\boldsymbol{K}$=4) of the same capacity can be delivered. Only full compartments are delivered to stations. Every petrol station $\boldsymbol{i}$ have a constant consumption $\boldsymbol{q_{ij}}$ for each fuel type $\boldsymbol{j}$, while intensity of consumption varies for different stations and different types of fuel. Petrol stations are equipped with underground tanks of known capacity $\boldsymbol{Q_{ij}}$ (one for each fuel type). Stations can be served only once during the day (observed interval). It is not allowed that inventory level for any of fuel types fall to zero. Vehicle fleet is of unlimited size, and therefore all deliveries can be realized. Furthermore, vehicles can not be outsourced and therefore fleet size is to be determined and minimized. Total inventory costs are assumed to be dependent on the sum of the stock level of each day in planning horizon, while transportation costs depend on travel distance and on the fleet size. Since there are maximally four compartments, in one route up to four petrol stations can be visited. That is, four possible routing strategies, visiting one, two, three or four petrol stations, can be used in fuel delivery.

3. MATHEMATICAL FORMULATION OF IRP

IRP model formulation presented here simultaneously minimizes inventory, routing and fleet size costs. Mathematical model that defines optimal time intervals and quantities for each delivery during the planning horizon is formulated as MIP and it comprises Inventory Costs (***IC***) and Transportation Costs (***TC***), where later includes Routing Costs (***RC***) and Fleet Costs (***FC***). Overall objective of the proposed model is to minimize total costs (1).

$$\textit{Minimize} \rightarrow \boldsymbol{IC + RC + FC} \tag{1}$$

$$\boldsymbol{IC} = \sum_{i}\sum_{j}\sum_{t}\left[\left(S^{0}{}_{ij}\text{-}t\cdot q_{ij} + \frac{q_{ij}}{2}\right) + \sum_{z=1}^{t}\sum_{k} x_{ijzk}\cdot d_k\right]\cdot c_{inv} \tag{2}$$

$$\boldsymbol{RC} = \sum_{t}\sum_{p=1}^{I}\sum_{q=p}^{I} y_{pqt}\cdot r_{pq} + \sum_{t}\sum_{p\in I}\sum_{q\in I\setminus\{p\}}\sum_{w\in I\setminus\{p,q\}} y_{pqwt}\cdot r_{pqw} + \sum_{t}\sum_{p\in I}\sum_{q\in I\setminus\{p\}}\sum_{w\in I\setminus\{p,q\}}\sum_{e\in I\setminus\{p,q,w\}} y_{pqwet}\cdot r_{pqwe} \tag{3}$$

$$\boldsymbol{FC} = t\cdot\left[F_a\cdot c_v + (F - F_a)\cdot c_m\right] \tag{4}$$

st.

$$S_{ij}^{0} + \sum_{t=1}^{z}\sum_{k} x_{ijtk}\cdot d_k - \sum_{t=1}^{z-1} q_{ij} \le Q_{ij} \quad \forall i \quad \forall j \quad \forall z \tag{5}$$

$$S_{ij}^{0} + \sum_{t=1}^{z}\sum_{k} x_{ijtk}\cdot d_k - \sum_{t=1}^{z} q_{ij} \ge q_{ij} \quad \forall i \quad \forall j \quad \forall z \tag{6}$$

$$\sum_{j}\sum_{k} x_{ijtk}\cdot k \le K \quad \forall i \quad \forall t \tag{7}$$

$$y_{pqt} \le \sum_{k}\sum_{j} x_{pjtk} \quad \forall p \quad \forall q \quad \forall t \tag{8}$$

$$y_{pqt} \le \sum_{k}\sum_{j} x_{qjtk} \quad \forall p \quad \forall q \quad \forall t \tag{9}$$

$$y_{pqwt} \le \sum_k \sum_j x_{pjtk} \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \tag{10}$$

$$y_{pqwt} \le \sum_k \sum_j x_{qjtk} \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \tag{11}$$

$$y_{pqwt} \le \sum_k \sum_j x_{wjtk} \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \tag{12}$$

$$y_{pqwet} \le \sum_k \sum_j x_{pjtk} \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \quad \forall e \in I \setminus \{p,q,w\} \tag{13}$$

$$y_{pqwet} \le \sum_k \sum_j x_{qjtk} \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \quad \forall e \in I \setminus \{p,q,w\} \tag{14}$$

$$y_{pqwet} \le \sum_k \sum_j x_{wjtk} \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \quad \forall e \in I \setminus \{p,q,w\} \tag{15}$$

$$y_{pqwet} \le \sum_k \sum_j x_{ejtk} \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \quad \forall e \in I \setminus \{p,q,w\} \tag{16}$$

$$H_{it} \le \sum_j \sum_k x_{ijtk} \quad \forall t \quad \forall i \tag{17}$$

$$H_{it} \ge \frac{1}{J} \cdot \sum_j \sum_k x_{ijtk} \quad \forall t \quad \forall i \tag{18}$$

$$\sum_p y_{ppt} + 2 \cdot \sum_t \sum_{p \in I} \sum_{q \in I \setminus \{p\}} y_{pqt} + 3 \cdot \sum_t \sum_{p \in I} \sum_{q \in I \setminus \{p\}} \sum_{w \in I \setminus \{p,q\}} y_{pqwt} + 4 \cdot \sum_t \sum_{p \in I} \sum_{q \in I \setminus \{p\}} \sum_{w \in I \setminus \{p,q\}} \sum_{e \in I \setminus \{p,q,w\}} y_{pqwet} = \sum_i H_{it} \quad \forall t \tag{19}$$

$$\sum_j \sum_k x_{pjtk} \cdot k \le K \cdot \left(2 - y_{ppt}\right) \quad \forall t \quad \forall p \tag{20}$$

$$\sum_j \sum_k \left(x_{pjtk} + x_{qjtk}\right) \cdot k \le K \cdot \left(2 - y_{pqt}\right) \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \tag{21}$$

$$\sum_j \sum_k \left(x_{pjtk} + x_{qjtk} + x_{wjtk}\right) \cdot k \le K \cdot \left(3 - 2 \cdot y_{pqwt}\right) \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \tag{22}$$

$$\sum_j \sum_k \left(x_{pjtk} + x_{qjtk} + x_{wjtk} + x_{ejtk}\right) \cdot k \le K \cdot \left(4 - 3 \cdot y_{pqwet}\right) \quad \forall t \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \quad \forall e \in I \setminus \{p,q,w\} \tag{23}$$

$$\begin{aligned} & y_{ppt} + \sum_t \sum_{q \in I \setminus \{p\}} \left(y_{pqt} + y_{qpt}\right) + \sum_t \sum_{q \in I \setminus \{p\}} \sum_{w \in I \setminus \{p,q\}} \left(y_{pqwt} + y_{qpwt} + y_{qwpt}\right) + \\ & + \sum_t \sum_{q \in I \setminus \{p\}} \sum_{w \in I \setminus \{p,q\}} \sum_{e \in I \setminus \{p,q,w\}} \left(y_{pqwet} + y_{qpwet} + y_{qwpet} + y_{qwept}\right) \le 1 \quad \forall t \quad \forall p \end{aligned} \tag{24}$$

$$\sum_{p=1}^{I} y_{ppt} + \sum_t \sum_{p \in I} \sum_{q \in I \setminus \{p\}} y_{pqt} + \sum_t \sum_{p \in I} \sum_{q \in I \setminus \{p\}} \sum_{w \in I \setminus \{p,q\}} y_{pqwt} + \sum_t \sum_{p \in I} \sum_{q \in I \setminus \{p\}} \sum_{w \in I \setminus \{p,q\}} \sum_{e \in I \setminus \{p,q,e\}} y_{pqwet} = F_t \quad \forall t \tag{25}$$

$$F \ge F_t \quad \forall t \tag{26}$$

$$F \ge F_a \quad \forall t \tag{27}$$

$$H_{it}, x_{ijtk}, y_{pqt}, y_{pqwt}, y_{pqwet} \in \{0,1\} \quad \forall i \quad \forall j \quad \forall p \quad \forall q \in I \setminus \{p\} \quad \forall w \in I \setminus \{p,q\} \quad \forall e \in I \setminus \{p,q,w\} \tag{28}$$

Where:

i,p,q,w,e	- petrol stations ($i, p, q, w, e \in \{1,2,\ldots,I\}$)
j	- fuel types ($j \in \{1,2,\ldots,J\}$)
t ,z	- time period or day in the planning horizon T ($t, z \in \{1,2,\ldots,T\}$)
k	- number of compartments
S^0_{ij}	- stock level of fuel type j at station i at the beginning of observed interval
d_k	- delivery quantities corresponding to the number of compartments
q_{ij}	- consumption of the fuel type j at station i
c_{inv}	- inventory carrying costs per day
Q_{ij}	- capacity of the underground reservoir for the fuel type j at station i
r_{pq}	- costs of visiting petrol stations p and q in one route of minimal length
r_{pqw}	- costs of visiting petrol stations p, q and w in one route of minimal length

r_{pqwe} - costs of visiting petrol stations p, q, w and e in one route of minimal length
F - required fleet size that consists of available vehicles and dummy vehicles
F_a - available fleet size
F_t - required number of vehicles per each time interval t (equals number of routes per each day)
c_v - fixed cost of available fleet size per vehicle per day
c_m - fixed cost of available fleet size for dummy vehicles per vehicle per day

$$x_{ijtk} = \begin{cases} 1 - \text{ if petrol station } \boldsymbol{i} \text{ is supplied with fuel type } \boldsymbol{j} \text{ in interval } \boldsymbol{t} \text{ with } \boldsymbol{k} \text{ compartments} \\ 0 - \text{ otherwise} \end{cases}$$

$$y_{ppt} = \begin{cases} 1 - \text{ if petrol station } \boldsymbol{p} \text{ is supplied in direct delivery at interval } \boldsymbol{t} \\ 0 - \text{ otherwise} \end{cases}$$

$$y_{pqt} = \begin{cases} 1 - \text{ if petrol station } \boldsymbol{p} \text{ and petrol station } \boldsymbol{q} \text{ are supplied in the same route at interval } \boldsymbol{t} \\ 0 - \text{ otherwise} \end{cases}$$

$$y_{pqwt} = \begin{cases} 1 - \text{ if petrol stations } \boldsymbol{p}, \boldsymbol{q} \text{ and } \boldsymbol{w} \text{ are supplied in the same route at interval } \boldsymbol{t} \\ 0 - \text{ otherwise} \end{cases}$$

$$y_{pqwet} = \begin{cases} 1 - \text{ if petrol stations } \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{w} \text{ and } \boldsymbol{e} \text{ are supplied in the same route at interval } \boldsymbol{t} \\ 0 - \text{ otherwise} \end{cases}$$

Inventory segment (2) of the objective function (1) tries to minimize total inventory costs of the fuel carried in all of petrol stations, during the observed planning horizon. Those costs are based on average daily inventory level at petrol station,. Routing segment (3) of th e objective function (1) tries to minimize total travel distance of all routes that a re used f or delivery du ring t he obse rved planning horizon by sol ving se t of assi gnment pr oblems. Routing co sts ar e calculated as sum of all travel costs incurred by deliveries in planning horizon. For each possible set of petrol stations travel costs are calculated for the route of minimal length, determined by enumeration. The third segment of objective function (1) represents fleet size costs (4) with the idea of minimizing maximal number of routes performed in one day, through out the planning horizon. Since each route represents one vehicle, when available fleet size can't provide all required deliveries dummy vehicles serve as reserve that may assure delivering all planned fuel quantities.

Constraints (5) limit maximal qu antity o f fuel up t o th e reservo ir cap acity in each of observed i ntervals, wh ile constraints (6) define minimal quantity of fuel in reservoirs which can meet demand in observed period. Constraints (7) prohibit multiple deliveries to th e same petrol station during the same day. If co nstraints (7) are omitted than different fuels m ay be delivered t o the sam e station, each in qua ntity lesser than four com partments, but in case when t otal number of compartments for all types of fuels delivered to the same station is greater then four then the station is to b e visited more than once in the same time period.. Constraints (8) to (16) defines that route can be performed only if all of petrol stations included in the route are to be supplied on that day. Inequalities (17) and (18) define stations that must be served. Pet rol st ation m ust be se rved i f a t l east one c ompartment of at least one fuel type s hould be delivered. Constraints (19) assure that number of stations served by all routes is equal to number of all stations that need to be served. Constraints from (20) to (23) allow that each route can have maximum four com partments in each day of the planning horizon. For example, if one station has delivery of three compartments, than this station can be served either with direct delivery or together with the station that needs delivery of only one compartment. Constraints (24) limit the number of routes visiting each of stations to only one. Equations (25) determine the number of routes per each d ay of the planning horizon. Constraints (26) and (27) determine required fleet size for delivery of all demanded quantities in the planning horizon. Constraints (28) define binary nature of variables.

4. HEURISTIC APPROACH

Heuristic proposed in this research tries to determine average delivery quantities per each day in planning horizon while equalizing vehicles' workload as a surrogate of fleet size, and while minimizing lengths of routes through appropriate matching of petrol station in single route. Workload equalizing is based on moving planned supplies one day earlier, but not later than the last day when supply must be realized in order to prohibit falling inventories to zero. Matching petrol stations satisfying during the interval t, tries to provide total travel distance shorter than in the case when each station is visited separately.

Steps of proposed heuristic are:

STEP 1: ***Determining plan of deliveries*** by solving the model (2), using constraints (5), (6), and (7)

STEP 2: ***Calculating average daily delivery quantity*** $\overline{\chi}$ in the observed planning horizon, T as

$$\overline{\chi} = \frac{1}{T}\sum_i \sum_j \sum_t \sum_k x_{ijtk} \cdot d_k \qquad (29)$$

STEP 3: ***Moving deliveries, one day earlier, one by one compartment***, until all daily delivery quantities are equalized. Compartments to be moved are chosen in accordance to the values of four criterions calculated for all of petrol stations, starting from the last day in the observed interval.

i) Value of the first criterion V_{ijt}^1 defines possibility of moving delivery of fuel type ***j*** to the station ***i*** from the day ***t***, to the day ***t-1***. If K_{it} is number of compartments to be delivered to the petrol station ***i***, during the day ***t***, then transfer is possible only when $K_{it} > 0 \wedge K_{it-1} < 4$ in which case $V_{ijt}^1 = 1$, otherwise $V_{ijt}^1 = 0$.

ii) Value of second criterion V_{ijt}^2 is determined in following way. When $K_{it}=1 \wedge 0<K_{it-1}<4$ then $V_{ijt}^2 = 2$ (one station will be deleted from observed day and its delivery quantity will be added to station that already exists in delivery plan for the day before). When $K_{it} > 1 \wedge K_{it-1} < 4$ then $V_{ijt}^2 = 1$, while otherwise $V_{ijt}^2 = 0$.

iii) Third criterion respects spatial grouping of stations which should decrease routing distance. To determine value of V_{ijt}^3 for every station ***i*** where $K_{it}>0$, do the following. Find the largest distance l_{ip} to the station ***p***, so that $K_{pt}>0$, and then find the shortest distance l_{ik}, so that $K_{k,t-1}>0$. Then calculate $V_{ijt}^3 = l_{ik} - l_{ip}$.

iv) The fourth criterion V_{ijt}^4 respects consumption rate where compartment of fuel type ***j*** for observed station that has lowest consumption is more preferable for transfer. This criterion takes reciprocally value of consumption rate, $V_{ijt}^4 = 1/q_{ij}$.

Based on eligibility (stations are primarily sorted by first criterion, then by second, then by third and finally by fourth), stations are transferred one by one until average delivery quantities are achieved (as long as there are available quantities). After each transfer eligibility is updated. Table 1 shows example of sorted compartments for transfer by four criterions of eligibility.

Table 1. Sorted compartments by eligibility for transfer to day before

Compartments	V_{ijt}^1	V_{ijt}^2	V_{ijt}^3	V_{ijt}^4
1	1 2 0			2
2	1 2 0			1
3	1 2 0			0.5
4 1		1	-17	1
5 1		1	-17	0.5
6 1		1	-28	1
7 1		0	-31	2
8* 0		2	-25	2
9* 0		1	-28	1
10* 0		0	-31	0.5

* - these compartments are not possible for transfer because there are already maximum number of compartment in delivery plan for observed station in the day before ($V_{ijt}^1=0$)

STEP 4: ***Matching petrol stations in single routes***. Matching of petrol stations is based on values of matching utilities.

i) Calculate utilities for all possible matchings and all permutations of its visit orders by applying (30) – (33) representing cases when no matching is considered (30), and when two (31), three (32) and four (33) petrol stations are matched in the single route ($p,q,w,e \in I$), Vehicle utilization ε={0.25, 0.5, 0.75, 1} is defined as a percentage of total number of four compartments are loaded in the route defined by the certain matching.

$$u_p = 2 \cdot K_{pt} \cdot l_{0p} - 2 \cdot \varepsilon \cdot l_{0p} \quad (30)$$

$$u_{pq} = 2 \cdot K_{pt} \cdot l_{0p} + 2 \cdot K_{qt} \cdot l_{0q} - \varepsilon \cdot \left(l_{0p} + l_{pq} + l_{qb}\right) \quad (31)$$

$$u_{pqw} = 2 \cdot K_{pt} \cdot l_{0p} + 2 \cdot K_{qt} \cdot l_{0q} + 2 \cdot K_{wt} \cdot l_{0w} - \varepsilon \cdot \left(l_{0p} + l_{pq} + l_{qw} + l_{0w}\right) \quad (32)$$

$$u_{pqwe} = 2 \cdot K_{pt} \cdot l_{0p} + 2 \cdot K_{qt} \cdot l_{0q} + 2 \cdot K_{wt} \cdot l_{0w} + 2K_{et} \cdot l_{0e} - \varepsilon \cdot \left(l_{0p} + l_{pq} + l_{qw} + l_{we} + l_{0e}\right) \quad (33)$$

ii) List utilities in decreasing order. Sequence of petrol stations with the largest utility gives the first route. The list is than updated by elimination of nodes still assigned to the route, and procedure continues until all routes are defined.

iii) Calculate number of routes R_t for each day of planning horizon, and calculate total costs T'_{IRP}

STEP 5: ***Solution improvement*** by transferring deliveries in order to reduce maximal number of routes during the day. Find the day $1 \le k \le T$, where $R_k = \max_{1 \le t \le T} (R_t)$. Then, find the day z < k, so that moving the delivery $K_{it} \rightarrow K_{jz}$ decreases maximal number of routes R_k. Calculate costs for the new delivery plan T''_{IRP}. If $T''_{IRP} < T'_{IRP}$, accept this delivery plan, otherwise reject. Repeat this step until all possible movements are done.

5. COMPUTATIONAL RESULTS

In this chapter we present computational results for MIP formulations and heuristics proposed. Testing the quality of solutions is carried out on the following numerical example:

- The number of petrol stations was I=10, each supplied with J=3 different fuel types during the planning horizon of T=5 days.
- Vehicles are equipped with tanks that have four compartments each with 8 t capacity.
- Stock levels of fuels S^0_{ij} at the beginning of observed interval are generated randomly between 1 and 8 t.
- Daily fuels consumptions q_{ij} are generated using discrete distribution, taking values of 0.5 t with probability of $p_{0.5}$=0.3, 1 t with probability of p_1=0.5 and 2 t with probability of p_2=0.2.
- Reservoirs capacities Q_{ij} have two possible values 20 t, or 30 t, which is randomly assigned to petrol stations and fuel types.
- Spatial coordinates of petrol stations are randomly generated between 2 and 97 km while depot is located at (0,0).
- Daily cost of carrying inventory is c_{inv}=0.82 €/t (equivalent to annual carrying inventory cost that is equal to 30% of stock value, if 1 lit of fuel has value of 1 €). Fixed cost of fleet size is c_v=200 € per vehicle per day for available vehicles and c_m=10000 € per dummy vehicle per day. Cost of one traveled kilometer is c_{km}=1 €/km.

For the purpose of evaluation the quality of proposed IRP model and heuristic approach we also solve models (34), and (35) separately. Coefficients of 0.0001 are used for the cases when more than one optimal solution existed.

Routing model: $Minimize \rightarrow 0.0001 * \boldsymbol{IC} + \boldsymbol{RC} + \boldsymbol{FC}$ (34)

Inventory model: $Minimize \rightarrow \boldsymbol{IC} + 0.0001 * (\boldsymbol{RC} + \boldsymbol{FC})$ (35)

Set of test examples consists of 10 replications where each is generated randomly. Test examples were solved using proposed IRP model, Routing model, Inventory model and heuristic approach. Comparison results for these methods are presented in Table 3. Table 2 presents input parameters for replication 4.

Table 2. Input parameters for replication ***4*** (q_{ij} [t], S^o_{ij} [t], Q_{ij} [t], (x,y) [km])

Station i	q_{i1}	q_{i2}	q_{i3}	S^o_{i1}	S^o_{i2}	S^o_{i3}	Q_{i1}	Q_{i2}	Q_{i3}	x	y
1	1	0.5	0.5	4	5	6	20	30	30	77	82
2	0.5	0.5	1	3	3	2	30	20	30	57	28
3	0.5	0.5	1	5	6	1	30	30	20	81	25
4	2	0.5	0.5	2	6	5	30	30	20	94	39
5	0.5	1	2	5	7	6	30	30	30	71	69
6	0.5	1	0.5	2	5	6	20	30	20	50	79
7	1	2	1	6	2	1	20	30	30	10	18
8	1	0.5	1	4	1	6	20	30	30	27	26
9	1	0.5	1	5	5	3	20	30	20	43	82
10	2	1	1	3	4	4	30	30	30	2	21

To test MIP model in reasonable time we have restrained model to triple assignment (maximum three stations can be served in a single route). This means that all y_{pqwe} variables are excluded from model together with utility (33) in routing step of heuristic.

Because MIP formulation includes dummy vehicles (see (4), (25), (26), and (27)) solutions are obtained through an iterative procedure where in the first is needed to define available number of vehicles (F_a). Through few iterations, number of dummy vehicles is changed until obtaining solution with minimal number of available vehicles without use of dummy vehicles which is optimal. This is why there are different fleet size costs in solutions of one problem instance solved by different models (for example, in the problem instance 10, fleet for IRP and Routing model has one vehicle, for Inventory model has three and for heuristics approach has two vehicles). Solution for the problem instance 4 is presented in Table 4.

MIP models were implemented through the CPLEX 12 on desktop PC with 2.0 GHz Dual Core processor with 2 GB of RAM memory. All input data needed for model implementation, as well as heuristics are implemented in Python 2.6.

6. CONCLUSION

Results in Table 3 shows that IRP model gives minimal total costs for all problem instances which is understandable because it considers both inventory and transport costs. Routing model gives 0.0017% higher total cost which is highly negligible difference. This fact can be explained by higher impact of transport costs on total IRP costs. On the other hand, if we would consider case of high valuable commodity, proposed IRP model would perform significantly better than Routing model because the impact of inventory costs would strengthen. Inventory model gives 75.4% higher total cost than optimal solution of IRP model. This confirms that inventory costs are not primal segment of IRP cost that can be used for optimization (small improvements of inventory costs led to much higher transportation costs).

Interesting conclusion can be drawn from the computational time needed for obtaining solution in case of MIP models. Average computational time for IRP model is 40 sec, for Routing model is 107 sec and for Inventory model is less than 1 sec although they use same formulation but with different impact of inventory and transport costs. This implies sensitivity of MIP models calculation time on weights of these segments in objective function.

Heuristic approach gives 16.5% greater total cost than optimal solution of IRP model and gives better solutions than the MIP Inventory model (both in total costs and computational time, which is less than 1s). Combination of computational time and quality of solutions obviously suggests use of heuristic on problems of real dimensions (more than 100 petrol stations with intensive deliveries) where MIP model cannot find optimal solution in reasonable time.

Proposed MIP problem is highly depended on relationship between inventory and routing costs. If one of these segments has significantly greater impact on total IRP cost than solving IRP can have same results as solution of only that segment. In many cases routing costs have significantly greater impact on total IRP costs (Lee et al. (2003)). Analysis of relationship between routing and inventory costs and its impact on IRP is one direction in further research. Other directions are related to application of other solution approaches, as well as metaheuristics, but also in further problem analysis in order to include additional requirements and limitations which exist in real systems.

ACKNOWLEDGMENT
This work was partially supported by Ministry of Science and technological Development Republic of Serbia, through the project TR 15018, for the period 2008-2010.

7. REFERENCES

Moin N.H. and Salhi S. (2007). Inventory Routing Problem: A Logistical Overview. *Journal of the Operational Research Society* 58: 1185-1194.

Cornilier F., Boctor F.F., Laporte G., Renaud J. (2008). A heuristic for the multi-period petrol station replenishment problem. *European Journal of Operational Research* 191: 295-305.

Webb I.R. and Larson L.C. (1995). Period and phase customer replenishment: a new approach to strategic inventory/routing problem. *Eur J Opl Res* 85: 132–148.

Viswanathan S. and Mathur K. (1997). Integrating routing and inventory decisions in one warehouse multiretailer, multiproduct distribution systems. *Mngt Sci* 43: 294–312.

Bertazzi L., Speranza M.G. and Ukovich W. (1997). Minimization of logistic costs with given frequencies. *Transport Res B* 31: 327–340.

Chi-Guhn Lee, Yavuz A. Bozer, Chelsea C. White III, 2003. A heuristic Approach and Properties of Optimal Solutions to the Dynamic IRP. Working paper, University of Toronto, Toronto, Ontario, Canada.

Ribeiro R. and Lourenco H.R. (2003). Inventory-routing model, for a multi-period problem with stochastic and deterministic demand. Working Paper, Department of Economics and Business and GRELIET, University Pompeu Fabra, Barcelona, Spain.

Bramel J and Simchi-Levi D (1997). The Logic of Logistics. Springer:New York.

Table 3. Comparison results of three MIP and heuristic model

REPLICATION	OPTIMAL SOLUTION															HEURISTIC SOLUTION					
	IRP MODEL					ROUTING MODEL					INVENTORY MODEL										
	INV.	TRANSPORT		**TOTAL (€)**	*TIME (sec)*	INV.	TRANSPORT		**TOTAL (€)**	*TIME (sec)*	INV.	TRANSPORT		**TOTAL (€)**	*TIME (sec)*	INV.	TRANSPORT		**TOTAL (€)**	*TIME (sec)*	
		Route	Fleet				Route	Fleet				Route	Fleet				Route	Fleet		*t`*	*t``*
1	689	822	2000	3511	*34*	755	787	2000	3542	*94*	617	1278	4000	5895	*0.34*	643	1149	2000	3792	*0.05*	*0.39*
2	600	634	1000	2234	*7*	627	609	1000	2236	*6*	568	778	3000	4346	*0.31*	620	877	1000	2497	*0.03*	*0.34*
3	716	682	1000	2398	*14*	716	682	1000	2398	*24*	631	1193	3000	4824	*0.33*	656	931	1000	2587	*0.03*	*0.36*
4	688	718	2000	3406	*9*	708	711	2000	3419	*14*	577	1184	3000	4761	*0.34*	602	1077	2000	3679	*0.05*	*0.39*
5	699	1096	2000	3795	*66*	699	1096	2000	3795	*177*	594	1719	4000	6313	*0.31*	620	1524	2000	4144	*0.06*	*0.37*
6	784	974	2000	3758	*169*	784	974	2000	3758	*651*	626	1661	4000	6287	*0.33*	645	1480	2000	4125	*0.08*	*0.41*
7	704	642	1000	2346	*53*	704	642	1000	2346	*65*	606	1130	3000	4736	*0.31*	671	982	1000	2653	*0.03*	*0.34*
8	671	727	1000	2398	*29*	671	727	1000	2398	*25*	579	1211	3000	4790	*0.33*	631	1015	2000	3646	*0.06*	*0.39*
9	601	713	1000	2314	*4*	601	713	1000	2314	*8*	548	1109	2000	3657	*0.31*	600	1056	1000	2656	*0.06*	*0.37*
10	631	1140	1000	2771	*10*	631	1140	1000	2771	*9*	592	1537	3000	5129	*0.31*	604	1321	2000	3925	*0.05*	*0.36*
AVG.	678	815	1400	**2893**	***40***	689	808	1400	**2898**	***107***	594	1280	3200	**5074**	***0.32***	629	1141	1600	**3370**	*0.05*	***0.37***

t` - solution time of heuristic

t`` - total time of solving Inventory model (first step of heuristic) plus solution time of heuristic

Table 4. – Delivery quantities and routes per each day of planning horizon in solution for the problem instance ***4*** of MIP formulations and heuristic approach

	Compartments that needs to be delivered					Delivery routes				
	1. day	2. day	3. day	4. day	5. day	1. day	2. day	3. day	4. day	5. day
IRP MODEL	7 4 4			4 0		[7,10], [2,3,4]	[1,5,8]	[6,9]	[7,10]	-
ROUTING MODEL	8 4 4			3 0		[7,10], [2,3,4]	[1,5,8]	[6,9]	[10]	-
INVENTORY MODEL	5 2 2			5 5		[3,4,7], [10]	[2,8]	[5,9]	[10], [6,8]	[4,6,9]
HEURISTICS	5 3 3			4 4		[3,4], [7,10]	[2,5,8]	[1,6,9]	[8,10]	[4,6,9], [7]

QUANTIFYING THE BULLWHIP EFFECT IN A SEASONAL SUPPLY CHAIN WITH STOCHASTIC LEAD TIME

Dong Won Cho[1] and Young Hae Lee[2*]

[1]Department of Industrial and Management Engineering
Hanyang University
Ansan,Gyeonggi-do, South Korea
e-mail: dwcmjcho@hanyang.ac.kr

[2*]Department of Industrial and Management Engineering
Hanyang University
Ansan,Gyeonggi-do, South Korea
e-mail: yhlee@hanyang.ac.kr

Abstract: In this study, we quantify the bullwhip effect in a seasonal two echelon supply chain with stochastic lead time. The bullwhip effect is the phenomenon of demand variability amplification when one moves away from the customer to the supplier in a supply chain. So, it is recognized that this effect poses very severe problems for a supply chain. The retailer faces external demand for a single product from end customers, where the underlying demand process is a seasonal autoregressive moving average, SARMA $(1,0)X(0,1)_s$ demand process. And the retailer employs a base stock periodic review policy to replenish its inventory from the upstream party every period using the minimum mean-square error forecasting technique. In order to quantify the bullwhip effect in a seasonal supply chain, we use the ratio of the variance of retailer's order quantities experienced by the supplier to the actual variance of demand quantities from the customers. So, we need to derive the results, such as variance of forecast error for the lead time demand and variance of order quantity etc. After the bullwhip effect based on these results was obtained, we investigate what parameter has an impact on the bullwhip effect and how large each parameter affects it. Specifically, we show that the seasonal phenomenon plays an important role in a seasonal supply chain.

Keywords: Supply chain management, Bullwhip effect, Seasonal autoregressive moving average process, Stochastic lead time

1. INTRODUCTION

Seasonal supply chain contains a seasonal phenomenon that has a main impact on material and information flows both in and between facilities, such as vendors, manufacturing and assembly plants, and distribution centers. The seasonal phenomenon of demand, which exists when a series fluctuates according to some seasonal factor, is a common occurrence in many supply chains. This phenomenon can especially intensify the bullwhip effect occurring severe problems in supply chains and eventually decrease supply chain profitability, the difference between the revenue generated from the final customer and the total cost across the supply chain.

It is essential that maximizing supply chain profitability should keep a supply chain everlastingly competitive advantage in fierce business environment. As basic approach to achieve this objective, each independent entity of a supply chain should be to maintain stable inventory levels to fulfill customer requests at a minimum cost. However, the bullwhip effect is recognized as the main one among various barriers internal and external to hinder this objective. The bullwhip effect is the phenomenon of the increasing amplification of variability in orders occurring within a supply chain the more one moves upstream. This amplification effect includes demand distortion described as the phenomenon where order to the suppliers tends to have larger variance than the sales to the buyer. The occurrence of the bullwhip effect in a supply chain poses severe problems such as lost revenues, inaccurate demand forecasts, low capacity utilizations, missed production schedules, ineffective transportations, excessive inventory investments, and poor customer service.

---
* Corresponding author

Forrester (1969) proves the evidence of the existence on the bullwhip effect. Sterman (1989) exhibites the same phenomenon through an experiment known as the beer game. In addition, Lee et al. (1997a, b) discoveres five main sources that may lead to the bullwhip effect, including demand signal processing, non-zero lead-time, order batching, rationaning game under shortage, and price fluctuations and promotions. They argue that eliminating its main causes may significantly reduce the bullwhip effect. In the concrete, the demand process, lead times, inventory policies, supply shortage and the forecasting techniques employed etc. have a significant influence on the bullwhip effect. Among these, forecasting techniques, inventory policies and to some extent replenishment lead times are controllable by supply chain members and hence can suitably be decided upon to mitigate the bullwhip effect. But demand process is uncontrollable because of external demand occurring at the customer. Demand process including a seasonal phenomenon is uncontrollable, too. So, it is important to understand the impact of the seasonal phenomenon on the bullwhip effect in a seasonal supply chain.

There are plentiful of studies in the area of the bullwhip effect including demand process, forecasting techniques, lead times and an ordering policy. Alwan *et al.* (2003) studied the bullwhip effect under an order-up-to policy applying the mean squared error optimal forecasting method to an AR(1) and investigated the stochastic nature of the ordering process for an incoming ARMA(1,1) using the same inventory policy and forecasting technique. Chen *et al.* (2000a, 2000b), Luong (2007), and Luong and Phien (2007) studied the bullwhip effect resulting from an order-up-to policy when simplified forecasting schemes including the moving average (MA) technique, the exponential weighted moving average (EWMA) technique, and the minimum mean square error (MMSE) technique were used for an autoregressive process such as AR(1) and AR(*p*). Zhang (2004) also investigates the impact of forecasting methods on the bullwhip effect in two-state supply chain with a first-order autoregressive demand process. In addition, they prove that increasing lead-time enhances the bullwhip effect regardless of the forecasting methods employed. And Kim *et al.* (2006) extends Chen *et al.* (2000a, 2000b) in that they include stochastic lead time and provide expressions for quantifying the bullwhip effect. Duc *et al.* (2008a) quantified the bullwhip effect for a two-stage supply chain in which the demand process followed ARMA(1,1), with an order-up-to policy with the MMSE method used at the retailer. They analytically investigated the effects of the autoregressive coefficient, the moving average coefficient, and constant lead time on the bullwhip effect. And Duc *et al.* (2008b) extends their study under the assumption of deterministic lead time into stochastic lead time.

For the case of a non-seasonal and non-stationary demand process, Graves (1999) studied the bullwhip effect resulting from a myopic base stock policy using an optimal forecasting technique for the autoregressive integrated moving average process, ARIMA(0,1,1). This study showed that the net inventory with a non-seasonal and non-stationary demand is more than that for an independent and identically distributed demand. To the best of our knowledge, there is a lack of previous studies concerned with the bullwhip effect considering both a seasonal and stationary (or non-stationary) demand process and stochastic lead time.

The objective of this paper is to quantify the bullwhip effect including stochastic lead time based on a replenishment situation that is similar to the one by Duc *et al.* (2008a, 2008b) in a seasonal two stage supply chain with one retailer and one supplier. In order to consider seasonal phenomenon, a seasonal autoregressive moving average process, SARMA(1,0) X $(0,1)_s$ (*s*: seasonal period) is employed as customer demand process.

The SARMA demand process integrates a non-seasonal ARMA and a seasonal ARMA demand process. In fact, the non-seasonal ARMA demand process often fits the time series of the demand process better than does a pure AR or a pure MA demand process (Pindyck and Rubinfeld, 1998). Likewise, this process can be applied to a seasonal ARMA demand process. In this work, we quantify the bullwhip effect in a seasonal supply chain considering stochastic lead time.

The organization of this paper is as follows. In Section 2, we describe the proposed seasonal supply chain model with the properties of the SARMA(1, 0) X $(0, 1)_s$ process. Section 3 presents the measure quantifying the bullwhip effect in a seasonal supply chain considering stochastic lead time. Finally, the conclusions of this study are presented in Section 4.

2. SUPPLY CHAIN MODEL

This section details modeling framework for quantifying the bullwhip effect in a seasonal supply chain considering stochastic lead time.

The following notations are used in this paper.

D_t customer demand quantity in period t

$\hat{D}_t$ forecast value of D_t

q_t ordered quantity in period t

S_t order-up-to level in period t

ϕ first-order autocorrelation coefficient

e_t forecast error in period t

Θ first-order seasonal moving average coefficient

δ constant of SARMA process

μ_d mean of SARMA process

σ_d^2 variance of the demand quantities

L_t order lead time in period t

μ_L mean of lead time

σ_L^2 variance of lead time

$D_t^{L_t}$ lead time demand, which is the total demand of period t through $t+L_t-1$

$\hat{D}_t^{L_t}$ lead time demand forecast that the retailer will face over L_t future periods

$\hat{\sigma}_t^{L_t}$ standard deviation of the lead time demand forecast error in period t

z normal z-score

s seasonal period ($s = 1,2,3,\cdots$.)

In this study, a measure of the bullwhip effect was developed for a seasonal two echelon supply chain with one retailer and one supplier. The supplier is assumed to have an unlimited supply. We assume that shortages are fully backlogged. The retailer employs a base stock policy, a simple order-up-to inventory policy, for replenishing its inventory. Within each period, events in the retailer occur in the following sequence: the retailer first reviews its inventory, determines its respective order quantities at the beginning of the period, receives the orders placed at the end of the lead time, and finally fills the demands from inventory in the next period.

In the ordering process of the retailer, the following assumptions are made: (1) demand forecasting for a SARMA (1, 0) X (0, 1)$_s$ process is performed with the MMSE technique; (2) there is a stochastic order lead time L_t for orders placed by the retailer in period t; (3) the lead time L_t is assumed to be stationary and independently and identically distributed (*iid*) with mean μ_L and variance σ_L^2; (4) the lead time and demand are independent; and (5) only one order is outstanding at any point of time.

Since it is assumed that the demand can be described by a SARMA (1, 0) X (0, 1)$_s$ process, we have

$$D_t = \delta + \phi D_{t-1} + e_t - \Theta e_{t-s} \qquad t = 1,2,3,\cdots. \tag{1}$$

where e_t is an *iid* normal process with a mean 0 and a variance σ_e^2. For the SARMA (1, 0) X (0, 1)$_s$ demand process to be stationary, the following relationship must be true:

$$E[D_t] = E[D_{t-1}] = \mu_d, \qquad t = 1,2,3,\cdots. \tag{2}$$

Hence, the following Eqs. (3) and (4) are valid.

$$\mu_d = \frac{\delta}{1-\phi}. \quad (3)$$

$$\sigma_d^2 = \frac{1-2\Theta\phi^s+\Theta^2}{1-\phi^2}\sigma_e^2 .$$
(4)

From Eqs. (3) and (4), it can be seen that in order for the SARMA (1, 0) X (0, 1)$_s$ process to be stationary, we should have $|\phi|<1$. Likewise, in order for the process to be invertible, we should have $|\Theta|<1$ (Box and Jenkins, 1976).

As mentioned earlier, the retailer employs a base stock policy. In the base stock policy, an order of quantity q_t is placed at the beginning of the period t so that the inventory position reaches a pre-specified order-up-to level. Therefore, the order quantity q_t can be given as

$$q_t = S_t - S_t + D_{t-1}. \quad (5)$$

Note that we allow q_t to be negative, in which case we assume, similarly to Lee *et al.* (1997b), that this excess inventory is returned without cost. In addition, the order-up-to level S_t can be determined by the lead time demand as

$$S_t = \hat{D}_t^L + z\hat{\sigma}_t^L, \quad (6)$$

where z is the normal z-score chosen to meet a desired service level. The service level is defined as the probability that demand is satisfied from the on-hand inventory. For *iid* demands from a normal distribution, the order-up-to level S_t from Eq. (6) is optimal under the assumption that there is no set up or fixed order cost (Nahmias, 1997). The optimal order-up-to level S_t can be determined from the inventory holding and shortage costs (Heyman and Sobel, 1984). However, since these costs cannot be accurately estimated in practice, the service level approach is often employed when the order-up-to level is to be determined. Given a service level, the order-up-to level S_t is determined by obtaining the lead time demand forecast $\hat{D}_t^{L_t}$ and the standard deviation of the lead time demand forecast error $\hat{\sigma}_t^{L_t}$. Note that the order-up-to level S_t is not fixed at the same value for all periods in order to accommodate the variances and correlation effects of the demand process. That is, it can be adaptively determined for each period separately after observing the recent demand.

3. THE BULLWHIPE EFFECT MEASURE

In this section, a measure of the bullwhip effect is developed. The bullwhip effect measure used in this paper is the ratio of variance in order quantity experienced by the supplier to the actual variance in the customer demand. This measure has been used in previous research (Chen *et al*., 2000a, b; Duc *et al*., 2008a, 2008b; Luong, 2007; Luong and Phien, 2007). Therefore, in order to develop the bullwhip effect measure in a seasonal supply chain, the lead-time demand forecast and the forecast error in the forecasting technique used for a SARMA (1, 0) X (0, 1)$_s$ process need to be derived first. Then, based on these results, the variance in the order quantity can be obtained.

3.1 Demand forecast and forecast error

Many previous studies have pointed out that simple forecasting techniques such as MA and ES may lead to specification errors, but the use of the MMSE does not (Alwan *et al*., 2003; Duc *et al.,* 2008a, b). Therefore, the smaller is the forecast error, the more accurate is the demand forecast and the smaller is the order variance. As a result, the MMSE technique used in this paper is expected to mitigate the bullwhip effect.

It is noted that the lead time demand at the retailer can be expressed as

$$D_t^{L_t} = D_t + D_{t-1} + \cdots + D_{t+L_t-1} = \sum_{i=0}^{L_t-1} D_{t+i}. \quad (7)$$

In addition, with the MMSE technique, the lead time demand forecast $\hat{D}_t^{L_t}$ that the retailer will face over L_t can be given as

$$\hat{D}_t^{L_t} = \hat{D}_t + \hat{D}_{t-1} + \cdots + \hat{D}_{t+L_t-1} = \sum_{i=0}^{L_t-1} \hat{D}_{t+i} \cdot \tag{8}$$

According to Box and Jenkins [2], for the SARMA (1, 0) X (0, 1)$_s$ process, $\hat{D}_t^{L_t}$ can be determined by Eq. (9).

$$\hat{D}_{t+i} = E\left[D_{t+i} \middle| D_{t-1}, D_{t-2}, \ldots\right], \tag{9}$$

from which an exact e xpression of $\hat{D}_{t+i}$ can be derived. The lead time demand at the retailer can be described using a recursive relation for a constant correlation coefficient. By applying the following equations,

$$D_{t+i} = \delta + \phi D_{t+i-1} + e_{t+i} - \Theta e_{t-s+i} \tag{10}$$

and

$$D_{t+i-1} = \delta + \phi D_{t+i-2} + e_{t+i-1} - \Theta e_{t-s+i-1} \cdot \tag{11}$$

We can obtain

$$D_{t+i} = (\phi+1)\delta + \phi^2 D_{t+i-2} + e_{t+i} + e_{t+i-1} - \Theta e_{t-s+i} - \Theta e_{t-s+i-1,} \cdot \tag{12}$$

By applying this procedure recursively,

$$D_{t+i} = \left(1-\phi^{i+1}\right)\mu_d + \phi^{i+1} D_{t-1} + \sum_{j=0}^{i} \phi^j e_{t+i-j} - \Theta \sum_{j=0}^{i} \phi^j e_{t-s+i-j} \cdot \tag{13}$$

Taking the conditional expectation from Eq. (9),

$$\hat{D}_{t+i} = \left(1-\phi^{i+1}\right)\mu_d + \phi^{i+1} D_{t-1} - \Theta \sum_{j=0}^{i} \phi^j E\left[e_{t-s+i-j} \middle| D_{t-1}, D_{t-2}, \ldots\right] \cdot \tag{14}$$

Note that the last term in Eq. (14) is a m odification factor for the forecast which accounts for the residual. Using Eq. (14), we can derive the following proposition.

Proposition 1. The MMSE-based lead time demand forecast depends on th e seasonal period s and can be e xpressed as

$$\hat{D}_t^{L_t} = \left(L_t - \frac{\left(1-\phi^{L_t}\right)\phi}{1-\phi}\right)\mu_d + \frac{\left(1-\phi^{L_t}\right)\phi}{1-\phi} D_{t-1} - \Theta \sum_{i=0}^{L_t-1} \frac{\left(1-\phi^{L_t-i}\right)}{1-\phi} E\left[e_{t-s+i} \middle| D_{t-1}, D_{t-2}, \ldots\right] \cdot \tag{15}$$

Proof. By substituting Eq. (14) into Eq. (8), the following relationship can easily be obtained.

$$\hat{D}_t^{L} = \sum_{i=0}^{L_t-1}\left(\left(1-\phi^{i+1}\right)\mu_d + \phi^{i+1} D_{t-1} - \Theta \sum_{j=0}^{i} \phi^j E\left[e_{t-s+i-j} \middle| D_{t-1}, D_{t-2}, \ldots\right]\right) = \left(L_t - \frac{\left(1-\phi^{L_t}\right)\phi}{1-\phi}\right)\mu_d + \frac{\left(1-\phi^{L_t}\right)\phi}{1-\phi} D_{t-1} - \Theta \sum_{i=0}^{L_t-1} \frac{\left(1-\phi^{L_t-i}\right)}{1-\phi} E\left[e_{t-s+i} \middle| D_{t-1}, D_{t-2}, \ldots\right] \tag{16}$$

From proposition 1, the lead time demand forecast depends on the seasonal period. As a result, we have

$$D_t^{L_t} - \hat{D}_t^{L_t} = \sum_{i=0}^{L_t-1} \frac{\left(1-\phi^{L_t-i}\right)}{1-\phi} e_{t+i} - \Theta \sum_{i=0}^{L_t-1} \frac{\left(1-\phi^{L_t-i}\right)}{1-\phi} e_{t-s+i} + \Theta \sum_{i=0}^{L_t-1} \frac{\left(1-\phi^{L_t-i}\right)}{1-\phi} E\left[e_{t-s+i} \middle| D_{t-1}, D_{t-2}, \ldots\right] \cdot \tag{17}$$

The following proposition gives the variance of the lead time forecast error based on MMSE.

Proposition 2. The variance of the lead time demand forecast error based on MMSE depends on the seasonal period s and can be given as

$$\left(\hat{\sigma}_t^{L_t}\right)^2 = \left(1+\Theta^2\right) E\left[\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right)^2\right]\sigma_e^2 + \Theta^2 E\left[\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right)^2 I_{s-i}\right]\sigma_e^2 - 2\Theta E\left[\sum_{i=0}^{L_t-1}\sum_{j=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right)\left(\frac{1-\phi^{L_t-j}}{1-\phi}\right)\left(r_{i,j-s} + r_{i,j} I_{s-j}\Theta\right)\right]\sigma_e^2 \tag{18}$$

where $I_k = \begin{cases} 1, k > 0 \\ 0, \ o.w \end{cases}$, $r_{i,j} = \begin{cases} 1, i = j \\ 0, \ o.w \end{cases}$.

Proof. Since $E\left[D_t^{L_t} - \hat{D}_t^{L_t} | L_t = L\right] = 0$ based on Eq. (17), we can obtain $E\left[D_t^{L_t} - \hat{D}_t^{L_t}\right] = E\left[E\left[D_t^{L_t} - \hat{D}_t^{L_t} | L_t = L\right]\right] = 0$. $E\left[\left(D_t^{L_t} - \hat{D}_t^{L_t}\right)^2 | L_t = L\right]$ can be expressed as

$$\begin{aligned}
E\left[\left(D_t^{L_t} - \hat{D}_t^{L_t}\right)^2 | L_t = L\right] &= E\left[\left(\sum_{i=0}^{L_t-1} \frac{\left(1-\phi^{L_t-i}\right)}{1-\phi} e_{t+i} - \Theta \sum_{i=0}^{L_t-1} \frac{\left(1-\phi^{L_t-i}\right)}{1-\phi} e_{t-s+i} + \Theta \sum_{i=0}^{L_t-1} \frac{\left(1-\phi^{L_t-i}\right)}{1-\phi} E\left[e_{t-s+i} | D_{t-1}, D_{t-2}, \ldots\right]\right)^2 | L_t = L\right] \\
&= E\left[\left(\sum_{i=0}^{L-1} \frac{\left(1-\phi^{L-i}\right)}{1-\phi} e_{t+i}\right)^2\right] + \Theta^2 E\left[\left(\sum_{i=0}^{L-1} \frac{\left(1-\phi^{L-i}\right)}{1-\phi} e_{t-s+i}\right)^2\right] + \Theta^2 E\left[\left(\sum_{i=0}^{L-1} \frac{\left(1-\phi^{L-i}\right)}{1-\phi} E\left[e_{t-s+i} | D_{t-1}, D_{t-2}, \ldots\right]\right)^2\right] \\
&\quad - 2\Theta E\left[\left(\sum_{i=0}^{L-1} \frac{\left(1-\phi^{L-i}\right)}{1-\phi} e_{t+i}\right)\left(\sum_{j=0}^{L-1} \frac{\left(1-\phi^{L-j}\right)}{1-\phi} e_{t-s+j}\right)\right] + 2\Theta E\left[\left(\sum_{i=0}^{L-1} \frac{\left(1-\phi^{L-i}\right)}{1-\phi} e_{t+i}\right)\left(\sum_{j=0}^{L-1} \frac{\left(1-\phi^{L-j}\right)}{1-\phi} E\left[e_{t-s+j} | D_{t-1}, D_{t-2}, \ldots\right]\right)\right] \\
&\quad - 2\Theta^2 E\left[\left(\sum_{i=0}^{L-1} \frac{\left(1-\phi^{L-i}\right)}{1-\phi} e_{t-s+i}\right)\left(\sum_{j=0}^{L-1} \frac{\left(1-\phi^{L-j}\right)}{1-\phi} E\left[e_{t-s+j} | D_{t-1}, D_{t-2}, \ldots\right]\right)\right]
\end{aligned} \tag{19}$$

Using $E\left[e_t e_{t+i}\right] = E\left[e_t\right]E\left[e_{t+i}\right] = 0$, $I_k = \begin{cases} 1, k > 0 \\ 0, \ o.w \end{cases}$, and $r_{i,j} = \begin{cases} 1, i = j \\ 0, \ o.w \end{cases}$, we can obtain Eq. (20).

$$E\left[\left(D_t^{L_t} - \hat{D}_t^{L_t}\right)^2 | L_t = L\right] = \left(1+\Theta^2\right) \sum_{i=0}^{L-1} \left(\frac{1-\phi^{L-i}}{1-\phi}\right)^2 \sigma_e^2 + \Theta^2 \sum_{i=0}^{L-1} \left(\frac{1-\phi^{L-i}}{1-\phi}\right)^2 I_{s-i} \sigma_e^2 - 2\Theta \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} \left(\frac{1-\phi^{L-i}}{1-\phi}\right)\left(\frac{1-\phi^{L-j}}{1-\phi}\right)\left(r_{i,j-s} + r_{i,j} I_{s-j} \Theta\right)\sigma_e^2 \tag{20}$$

where $I_k = \begin{cases} 1, k > 0 \\ 0, \ o.w \end{cases}$, $r_{i,j} = \begin{cases} 1, i = j \\ 0, \ o.w \end{cases}$.

So, $E\left[\left(D_t^{L_t} - \hat{D}_t^{L_t}\right)^2\right] = E\left[E\left[\left(D_t^{L_t} - \hat{D}_t^{L_t}\right)^2 | L_t = L\right]\right]$ can be expressed as

$$E\left[\left(D_t^{L_t} - \hat{D}_t^{L_t}\right)^2\right] = \left(1+\Theta^2\right) E\left[\sum_{i=0}^{L-1} \left(\frac{1-\phi^{L-i}}{1-\phi}\right)^2\right] \sigma_e^2 + \Theta^2 E\left[\sum_{i=0}^{L-1} \left(\frac{1-\phi^{L-i}}{1-\phi}\right)^2 I_{s-i}\right] \sigma_e^2 - 2\Theta E\left[\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} \left(\frac{1-\phi^{L-i}}{1-\phi}\right)\left(\frac{1-\phi^{L-j}}{1-\phi}\right)\left(r_{i,j-s} + r_{i,j} I_{s-j} \Theta\right)\right]\sigma_e^2 \tag{21}$$

where $I_k = \begin{cases} 1, k > 0 \\ 0, \ o.w \end{cases}$, $r_{i,j} = \begin{cases} 1, i = j \\ 0, \ o.w \end{cases}$.

Finally, $\left(\hat{\sigma}_t^{L_t}\right)^2 = VAR\left(D_t^{L_t} - \hat{D}_t^{L_t}\right)$ can be expressed as

$$\left(\hat{\sigma}_t^{L_t}\right)^2 = VAR\left(D_t^{L_t} - \hat{D}_t^{L_t}\right) = E\left[\left(D_t^{L_t} - \hat{D}_t^{L_t}\right)^2\right] - \left\{E\left[D_t^{L_t} - \hat{D}_t^{L_t}\right]\right\}^2 = E\left[\left(D_t^{L_t} - \hat{D}_t^{L_t}\right)^2\right]. \tag{22}$$

From Eq. (22), it is shown that the variance of the lead time demand forecast error depends on the seasonal period.

3.2 Measure of the bullwhip effect

Using Eqs. (5), (6), and (15), the ordered quantity of the retailer can be described in terms of the customer demand process and can be express as

$$\begin{aligned}
q_t &= S_t - S_t + D_{t-1} = \left(\hat{D}_t^L - \hat{D}_{t-1}^L\right) + z\left(\hat{\sigma}_t^L - \hat{\sigma}_{t-1}^L\right) + D_{t-1} = \left(\hat{D}_t^L - \hat{D}_{t-1}^L\right) + D_{t-1} \\
&= \left(L_t - L_{t-1} + \frac{\left(\phi^{L_t} - \phi^{L_{t-1}}\right)\phi}{1-\phi}\right)\mu_d + \frac{\left(1-\phi^{L_t+1}\right)}{1-\phi} D_{t-1} - \frac{\left(1-\phi^{L_{t-1}}\right)}{1-\phi} D_{t-2} \\
&\quad - \sum_{i=0}^{L_t-1} \frac{\left(1-\phi^{L_t-i}\right)\Theta}{1-\phi} E\left[e_{t-s+i} | D_{t-1}, D_{t-2}, \ldots\right] + \sum_{i=0}^{L_{t-1}-1} \frac{\left(1-\phi^{L_{t-1}-i}\right)\Theta}{1-\phi} E\left[e_{t-s+i-1} | D_{t-2}, D_{t-3}, \ldots\right]
\end{aligned} \tag{23}$$

Therefore, the following proposition can be derived.

Proposition 3. $B\left(L_t, s, \Theta, \phi\right)$ depends on the seasonal period s and can be given as

$$s = 1,\ B\left(L_t, s, \Theta, \phi\right) = \frac{2\sigma_L^2 \mu_d^2}{\sigma_d^2} + \frac{E\left(\left(1-\phi\right)^2\left(1-\Theta\right)\left(1-\Theta-2\phi^{L_t}\left(\phi-\Theta\right)\right) + \left(\phi-\Theta\right)^2\left(\phi^{2L_t} + \phi^{2L_{t-1}} - 2\phi^{L_t+L_{t-1}+1}\right)\right)}{\left(1-2\Theta\phi+\Theta^2\right)\left(1-\phi\right)^2}. \tag{24}$$

$$\begin{aligned} s \geq 2,\ B(L_t, s, \Theta, \phi) &= \frac{2\sigma_L^2 \mu_d^2}{\sigma_d^2} + \frac{E\begin{bmatrix} \left(1-2\phi^{L_t+1}+\phi^{2L_t+2}+\phi^2\left(1-2\phi^{L_t-1}+\phi^{2L_t}\right)\right)\left(1-2\Theta\phi^s+\Theta^2\right) \\ -2\phi^2\left(1-\phi^{L_t-1}-\phi^{L_t+1}+\phi^{L_{t-1}+L_t+1}\right)\left(1-\Theta\phi^s+\Theta^2-\Theta\phi^{s-2}\right) \end{bmatrix}}{\left(1-2\Theta\phi^s+\Theta^2\right)(1-\phi)^2} \\ &+ \Theta^2\left(\frac{1-\phi^2}{1-2\Theta\phi^s+\Theta^2}\right)E\left[\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right)^2 I_{s-i} + \sum_{i=0}^{L_{t-1}-1}\left(\frac{1-\phi^{L_{t-1}-i}}{1-\phi}\right)^2 I_{s-i}\right] \\ &- 2\Theta\left(\frac{1-\phi^2}{1-2\Theta\phi^s+\Theta^2}\right)E\left[\left(\frac{1-\phi^{L_t+1}}{1-\phi}\right)\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right)\phi^{s-i-1} I_{s-i}\right] \\ &+ 2\Theta\left(\frac{1-\phi^2}{1-2\Theta\phi^s+\Theta^2}\right)E\left[\left(\frac{1-\phi^{L_t+1}}{1-\phi}\right)\left(\sum_{i=0}^{L_{t-1}-1}\left(\frac{1-\phi^{L_{t-1}-i}}{1-\phi}\right)\phi^{s-i} I_{s-i} - \Theta\left(\frac{1-\phi^{L_t-1}}{1-\phi}\right) I_s\right)\right] \\ &+ 2\Theta\left(\frac{1-\phi^2}{1-2\Theta\phi^s+\Theta^2}\right)E\left[\left(\frac{\left(1-\phi^{L_t-1}\right)\phi}{1-\phi}\right)\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right)\phi^{s-i-2} I_{s-i}\right] \\ &- 2\Theta\left(\frac{1-\phi^2}{1-2\Theta\phi^s+\Theta^2}\right)E\left[\left(\frac{\left(1-\phi^{L_t-1}\right)\phi}{1-\phi}\right)\sum_{i=0}^{L_{t-1}-1}\left(\frac{1-\phi^{L_{t-1}-i}}{1-\phi}\right)\phi^{s-i-1} I_{s-i}\right] \\ &- 2\Theta^2\left(\frac{1-\phi^2}{1-2\Theta\phi^s+\Theta^2}\right)E\left[\left(\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right) I_{s-i}\right)\left(\sum_{j=0}^{L_{t-1}-1}\left(\frac{1-\phi^{L_{t-1}-j}}{1-\phi}\right) I_{s-j}\right) r_{i,j-1}\right] \end{aligned} \tag{25}$$

where $I_k = \begin{cases} 1, k>0 \\ 0,\ o.w \end{cases}$, $r_{i,j} = \begin{cases} 1, i=j \\ 0,\ o.w \end{cases}$.

Proof. Since $E[q_t|L_t, L_{t-1}] = (L_t - L_{t-1} + 1)\mu_d$ based on Eq. (5), we can obtain $E[q_t] = E[E[q_t|L_t, L_{t-1}]] = \mu_d$. And using Eq. (5), $E[q_t^2|L_t, L_{t-1}]$ can be expressed as

$$\begin{aligned} E\left[q_t^2|L_t, L_{t-1}\right] &= (L_t - L_{t-1}+1)^2\mu_d^2 + \frac{E\begin{bmatrix} \left(1-2\phi^{L_t+1}+\phi^{2L_t+2}+\phi^2\left(1-2\phi^{L_t-1}+\phi^{2L_t}\right)\right)\left(1-2\Theta\phi^s+\Theta^2\right) \\ -2\phi^2\left(1-\phi^{L_t-1}-\phi^{L_t+1}+\phi^{L_{t-1}+L_t+1}\right)\left(1-\Theta\phi^s+\Theta^2-\Theta\phi^{s-2}\right) \end{bmatrix}}{\left(1-2\Theta\phi^s+\Theta^2\right)(1-\phi)^2} \\ &+ \Theta^2\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right)^2 I_{s-i}\sigma_e^2 + \Theta^2\sum_{i=0}^{L_{t-1}-1}\left(\frac{1-\phi^{L_{t-1}-i}}{1-\phi}\right)^2 I_{s-i}\sigma_e^2 - 2\Theta\left(\frac{1-\phi^{L_t+1}}{1-\phi}\right)\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right)\phi^{s-i-1} I_{s-i}\sigma_e^2 \\ &+ 2\Theta\left(\frac{1-\phi^{L_t+1}}{1-\phi}\right)\left(\sum_{i=0}^{L_{t-1}-1}\left(\frac{1-\phi^{L_{t-1}-i}}{1-\phi}\right)\phi^{s-i} I_{s-i} - \Theta\left(\frac{1-\phi^{L_t-1}}{1-\phi}\right) I_s\right)\sigma_e^2 + 2\Theta\left(\frac{\left(1-\phi^{L_t-1}\right)\phi}{1-\phi}\right)\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right)\phi^{s-i-2} I_{s-i}\sigma_e^2 (s\geq 2) \\ &- 2\Theta\left(\frac{\left(1-\phi^{L_t-1}\right)\phi}{1-\phi}\right)\sum_{i=0}^{L_{t-1}-1}\left(\frac{1-\phi^{L_{t-1}-i}}{1-\phi}\right)\phi^{s-i-1} I_{s-i}\sigma_e^2 - 2\Theta^2\left(\sum_{i=0}^{L_t-1}\left(\frac{1-\phi^{L_t-i}}{1-\phi}\right) I_{s-i}\right)\left(\sum_{j=0}^{L_{t-1}-1}\left(\frac{1-\phi^{L_{t-1}-j}}{1-\phi}\right) I_{s-j}\right) r_{i,j-1}\sigma_e^2 \end{aligned} \tag{26}$$

where $I_k = \begin{cases} 1, k>0 \\ 0,\ o.w \end{cases}$, $r_{i,j} = \begin{cases} 1, i=j \\ 0,\ o.w \end{cases}$.

Using $E[L_t^2] = E[L_{t-1}^2] = \sigma_L^2 + \mu_L^2$, $E[L_t L_{t-1}] = \mu_L^2$, and $E[q_t^2] = E[E[q_t^2|L_t, L_{t-1}]]$, we can obtain Eq. (24) and Eq. (25).

4. CONCLUSIONS

In th is paper, we attem pted to quantify t he bullwhip eff ect in a sea sonal two-ec helon supply chain c onsidering stochastic lead time and to identify the parameters which have significant impacts on the bullwhip effect. Our analytical results show t hat the aut oregressive coefficient, lead time, seasonal moving c oefficient, a nd the seas onal period influences t he bullwhip e ffect. These fi ndings p rovide m anagerial i nsights for t he ret ailer an d su pplier t o re duce t he bullwhip effect. Based on this study, further research will illustrate the magnitude of the impact of the parameters on the bullwhip effect using numerical experiments.

ACKNOWLEDGEMENTS

This r esearch was supported b y Ba sic S cience Re search Program th rough t he National Research Foundation of Ko rea (NRF) funded by the Ministry of Education, Science and Technology (2010-0016024)

5. REFERENCES

Alwan, L. C., Liu, J. and Yao, D. Q. (2003). Stochastic Characterization of Upstream Demand Processes in a Supply Chain, *IIE Transactions, 35:* 207-219.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*. Revised Edition, Holden-Day, California.

Chen, F., Drezner, Z. Ryan, J. K. and Simchi-Levi, D. (2000a). Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times, and Information, *Management Science, 46:* 436-443.

Chen, F., Ryan, J. K. and Simchi-Levi, D. (2000b). The Impact of Exponential Smoothing Forecasts on the Bullwhip Effect, *Naval Research Logistics*, 47: 269-286.

Duc, T. T. H., Luong, H. T. and Kim, Y.-D. (2008a). A Measure of Bullwhip Effect in Supply Chains with a Mixed Autoregressive-moving Average Demand Process, *European Journal of Operational Research*, 187: 243-256.

Duc, T. T. H., Luong, H. T. and Kim, Y.-D. (2008b). A Measure of the Bullwhip Effect in Supply Chains with stochastic lead time, *The International Journal of Advanced Manufacturing Technology, 38:* 1201-1212.

Forrester, J. W. (1969). *Industrial Dynamics*, MIT Press, Cambridge.

Graves, S. C. (1999), A Single-Item Inventory Model for a Nonstationary Demand Process, *Manufacturing & Service Operations Management, 1:* 50-61.

Heyman, D., and Sobel, M. (1984). *Stochastic Models in Operations Research*, McGraw-Hill, New York.

Kim, J. G., Chatfield, D., Harrison, T. P. and Hayya, J. C. (2006). Quantifying the Bullwhip Effect in a Supply Chain with Stochastic Lead Time, *European Journal of Operational Research, 173:* 617-636.

Lee, H. L., Padmanabhan, V. and Whang, S. G. (1997a). The Bullwhip Effect in Supply Chain, *Sloan Management Review, 38:* 93-102.

Lee, H. L., Padmanabhan, V. and Whang, S. G. (1997b). Information Distortion in a Supply Chain: The Bullwhip Effect, *Management Science, 43:* 546-558.

Luong, H. T. (2007). Measure of Bullwhip Effect in Supply Chains with Autoregressive Demand Process, *European Journal of Operational Research, 180:* 1086-1097.

Luong, H. T. and Phien, N. H (2007). Measure of Bullwhip Effect in Supply Chains: The case of high order autoregressive demand process, *European Journal of Operational Research, 183:* 197-209.

Nahmias, S. (1997), *Production and Operations Analysis*, 3rd ed., Irwin McGraw-Hill, New York.

Sterman, J. D. (1989), Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment, *Management Science*, 35: 321-339.

Zhang, X. (2004). The Impact of Forecasting Methods on the Bullwhip Effect, *International Journal of Production Economics, 88:* 15-27.

Session A2: Inventory and Production

·Day1: Sep. 15 (Wed.)

·Time: 09:00 - 10:20

·Chair: Roel G. van Anholt

·Room: Azalea & Lilac, 5F

AN INTEGRATIVE ONLINE ATM FORECASTING AND REPLENISHMENT MODEL WITH A TARGET FILL RATE

Roel G. van Anholt[1], Iris F.A. Vis[2]

Department of Information Systems and Logistics
Faculty of Economics and Business Administration
VU University Amsterdam
De Boelelaan 1105, 1081 HV, Amsterdam, The Netherlands
[1]e-mail: ranholt@feweb.vu.nl
[2]e-mail: ivis@feweb.vu.nl

Abstract: This paper presents an integrative analytical online forecasting and replenishment model for Automated Teller Machines (ATMs) to simultaneously calculate efficient ordering quantities and -moments. While dealing with a target fill rate, a trade-off is being made between ordering cost (i.e., costs for order preparation, sorting cash, transport and actual replenishment) and inventory holding cost (i.e., missed interest yield for storing cash in ATMs). The model considers a novel combination of characteristics, which are hourly demand forecasts with a trend overlay, variable lead times, capacitated inventory at ATMs and real time information on stock levels. We compare the proposed forecasting method decomposition with an Artificial Neural Network approach (ANN) and we perform exhaustive numerical simulation experiments to validate and to test the robustness of the integrative model. The results show that decomposition outperforms the ANN approach and that the integrative model is capable of performing well under a wide range of circumstances.

1. INTRODUCTION

Automated Teller Machines (ATM) are computerized telecommunication devices which provide financial institution's customers a method of financial transactions in a public space without the need for a human clerk (Simutis *et al.*, 2008). ATMs are part of a so-called 'cash supply chain', 'cash chain' or 'currency chain', generally consisting of a Central Bank, mints/banknote printers, a distribution network, commercial banks, public customers and businesses. In a cash chain there is a forward and backward motion of coins and notes with the purpose of distributing cash to the public and keeping the cash in circulation 'fit' (i.e., valid for circulation). The return flow of cash is serving the purpose of removing all un-fit cash (especially notes) back to the Central Bank. The cash chain can be classified as a closed-loop supply chain because ideally no cash leaves circulation over time. ATMs are supplied/replenished by the distribution network and public or business customers demand/withdraw cash. Among a wide variety of ATM types currently deployed globally, the classical ATM with the ability to only dispense banknotes is still being used most frequently. The main focus of this paper is on classical ATMs, although the overall structure of the solution method can be applied for a broader class of ATMs

Cash supply chain processes can be described as follows: ATM replenishments are ordered by a financial institution (e.g., a commercial bank or ATM owner). At cash centers – where cash is counted, sorted and checked for fit/unfit – the orders are prepared and transferred to a logistic service provider, also called the replenishing company. Usually the cash center requires the order to be placed before a specific moment in time to be able to prepare and transfer the order on the same day. The replenishing company stores the orders of cash temporarily and replenishes the ATMs within a specific lead time which is usually one up to three days. The lead time is stated in a contract called a service level agreement, which is signed by both the financial institution and the replenishing company.

A complex decision problem for financial institutions is to decide when to place replenishment orders and what the order quantity should be to on one hand overcome stock outs and on the other hand to minimize ordering and inventory holding cost. Knowledge of the most likely moment that an ATM runs out of cash is required to effectively deal with order lead times. Demand forecasting is crucial in this context as also observed by Simutis *et al.* (2007). Consequently, we design an accurate and reliable demand forecast method and a replenishment policy in an integrative way.

The objective of this paper is to present a novel dynamic solution approach that considers both ATM demand forecasting and ATM replenishment in an integrative way. Our main goal is to design a method that is understandable and that is ready to be implemented in practice right away. In achieving this goal, our research consisted of three steps, namely analyzing current practices, designing the model and validating the solution approach. To effectively deal with the first step, we focused in 2009 on methods and policies being used by major banks in The Netherlands and analyzed their performance. Although we do not report on the exact outcomes of this project, we will present general insights, show how we used the knowledge in designing the model and use part of the data acquired in the validation process. In Section 2 we discuss relevant literature to show the added value of our model. Section 3 encompasses the model description. In sections 4 and 5 we validate the proposed forecasting method and perform exhaustive numerical simulation experiments to test the robustness of the integrative model. Section 6 presents conclusions and further research issues.

2. LITERATURE OVERVIEW

Standard forecasting techniques, such as extrapolation, quantitative analogies and rule-based forecasting (Armstrong and Green, 2005) can be used to perform demand forecasting for ATMs. More recently, techniques as Artificial neural networks (ANN) and support vector regression analyses (SVR) have been introduced to forecast demand for cash money in ATMs(Simutis *et al.*, 2008). Although the SVR analysis is considered as a promising new technique to deal with demand forecasting, a flexible ANN – in which the generalization properties were improved using a special adaptive regularization term – outperforms the SVR approach by far according to Simutis *et al.* (2008).

In practice, the decision making company – such as a bank or ATM owner – usually acquires a commercial decision support system to perform ATM inventory management (Simutis *et al.*, 2008). Examples are ProCash Analyzer (Wincor Nixdorf), OptiCash (Transoft Internaltional), iCom (Carreker Corporation) and MorphisCM (Morphis Inc). Unfortunately the software vendors are not willing to share any detailed knowledge about the techniques used in their software, which is observed by Castro (2009) and by research performed at the VU University Amsterdam as well. According to Simutis *et al.* (2008), the commercial decision support systems make use of simple linear regression models to forecast future cash withdrawals. According to these authors, linear regression models have two major drawbacks when compared to new approaches like SVR and ANN:

1. Each ATM requires a different regression model with specific ATM-related coefficients. Performing the same method for multiple individual ATMs simultaneously is a difficult task.
2. The parameters of the model are determined in the system implementation stage and are left unchanged during the operation phase. Since the business environment is changing continuously in practice, the parameters should be altered accordingly.

To ensure a quick implementation of our method in practice, we choose not to design an SVR or ANN based approach, but instead we intend to improve the standard regression models with time-based coefficients. We show by means of experiments that we overcome the drawbacks mentioned (see Section 4). We suggest to use decomposition as an important part of the model. Decomposition consists of multiple sequential steps in which many calculations take place. Simple calculations are performed in each iteration to obtain a forecast per time unit by using up-to-date historical demand figures. According to Al-Saba and El-Amin (1999) decomposition is comparable to other time-series models like Box-Jenkins, simple exponential smoothing and Hot-Winters. Al-Saba and El-Amin (1999) demonstrate that an ANN approach outperforms other regression methods. However, they do not include decomposition in this test. We will perform these tests for decomposition and show its added value in Section 4.

The resulting demand estimation is required for finding the reorder level (i.e., the threshold on-hand inventory level by which an order must be placed) and the order quantity. Inventory management literature finds its origin in 1913 when Harris (1913) developed the economic order quantity model. Up till now, a wide variety of models and policies have been developed, each addressing a slightly different problem. Although the mature state of this type of research, new studies are performed these days to keep up with changing demands from practice. For example, more and more research is being performed to comply with a desired customer service level (e.g., a target fill-rate), instead of giving a number one priority to minimizing cost. The problem of ATM inventory management can be classified as a lost-sales problem, because transactions are lost in the case of stock outs; customers will travel to a nearby ATM to withdraw cash instead of wait until the ATM gets replenished. Next to that, ATM inventory management is usually confronted with a fixed ordering cost; because the cash transporting and replenishing company is often a third party and charges these costs to the decision making company (e.g., a bank or ATM-owner). According to a recent literature review (Bijvank and Vis, 2010) adequate solution approaches for these type of problems – lost-sales and fixed ordering cost – are difficult to obtain, especially when considering a positive lead time and a target fill rate. Applicable exact approaches for this type of problem are policy iteration algorithms and numerical search procedures. Hill and Johansen (2006) propose a policy iteration algorithm to find optimal order quantities while considering a zero lead time. A periodic review policy with fixed order quantities (i.e., (R,s,Q) policy) with a cost minimization objective is developed by Johansen and Hill (2000). Bijvank *et al.* (2010) develop mathematical models for fixed order size policies and order-up-to policies while no assumptions on the number of outstanding orders are made. Tijms and Groenevelt (1984) propose a model considering a target fill rate with an order-up-to policy. None of these studies consider non-stationary periodic reviews (i.e., review moments only on specific moments in time). In an ATM inventory management setting, inventory cannot be replenished on all weekdays due to the working days of the replenishing company. We included this realistic aspect in the our model which has far-reaching consequences for finding efficient values for the reorder levels and order quantities as we will demonstrate in the design of our solution approach in Section 3.2.

In conclusion, we propose an integrative demand forecasting and replenishment model. We introduce transaction data as the main input for our model. So far, to our knowledge, decision support systems for cash management deal with daily demand figures. Transaction data contain information of each customer withdrawal, and allow for more accurate forecasting resulting in hourly forecasts of demand. A drawback might be the magnitude of the required dataset which might result in more complex and time-consuming modeling and processing. The results in this paper demonstrate how to overcome this drawback. We suggest to initiate replenishment decisions by means of continuously forecasting demand. So far, replenishments are triggered on fixed moments in time when the inventory level drops below the static reorder level. Our approach is capable of continuously recalculating the reorder level by taking into account a real time demand forecast. As a result, continuous monitoring of both forecasts and stock levels, enables to detect stock outs in an early stage or even before they happen. Emergency supplies or in-house transshipments to the specific ATM can be initiated based on this information. Regular replenishments normally need to be placed at a specific moment to ensure strict lead times. We will show in Section 3 that the design of the integrative component in our method allows for both options.

3. INTEGRATIVE ATM DEMAND FORECASTING AND REPLENISHMENT MODEL

This section describes the integrative demand forecasting and currency replenishment model. Section 3.1 discusses the problem description and Section 3.2 the model logic.

3.1 Problem description

The general objective is to minimize long-run average cost while considering a target fill rate, by calculating when to order and how much to order (i.e., the ordering quantity). The obtained demand forecast, order levels and order quantities comprise the main control variables. The actual fill rate together with the resulting replenishment- and interest costs are the main performance indicators. Many realistic constraints are taken into account to make sure the model can be deployed in practice. In designing the model, we assumed ATMs to be independent from one another, that means the model performs inventory management for individual ATMs with a known capacity only. As explained in Section 2, we design a model that both allows for continuous monitoring based on real-time demand forecasts and for non-stationary periodic order moments (i.e., orders can be placed only at specific given moments in time, e.g., on working days at 11:00 AM). The model assumes lost-sales. The order quantity is variable and is calculated as a real currency (e.g., € 163,746.23). Fixed costs are involved for each replenishment and the interest rate determines the on-hand inventory holding costs. Instead of assuming penalty costs for stock outs, we deal with a target fill rate. No costs are charged for having excess cash at replenishments, however the model tries to reduce it by performing several checks. Using transaction data as input, the model calculates the arrival rate of consumers (normally distributed) and the average amount people withdraw per hour (empirical distribution with hourly changing parameters). Days with unusual demand patterns (e.g., holidays) are considered separately. We assume stochastic lead times and independent replenishment orders in the case of multiple outstanding orders. At each replenishment the ordering quantity is added to the remaining stock. We do not consider ‘cassette swapping’ by which all the remaining stock is returned to the cash center and replaced by new cash money. The model input parameters are summarized in Table 1.

3.2 Model logic

We distinguish between five steps in our model which are depicted in Figure 1. The approach starts with the preparation of the historical demand data which is used in the second stage to filter out time-related patterns. By calculating a moving average and overlaying this average with the time-related patterns, we derive a demand forecast. The first two steps – data preparation and filtering out time-related patterns – should be performed daily or weekly. The moving average should be updated continuously with newly obtained historical demand figures. The demand forecast (which is calculated with the moving average) is used for calculating both the reorder level and the order quantity which respectively answer the questions ‘when to order’ and ‘how much to order’. The question ‘when to order’ can be answered either on specific moments in time or continuously. The demand forecast can be calculated up to infinity, but one should keep in mind that forecasting over longer periods will be based on the moving average calculated at t_0. So forecasts will be less accurate for longer periods. As discussed, the model is capable of checking continuously or multiple times a day whether an order should be placed.

Table 1. Model input parameters

Transaction data	A dataset consisting of the time and transaction amount of all individual withdrawals of the ATM under consideration over a period of time. At least from previous three months but preferably from last 3 years.
Holidays	A list of all holidays and similar days which show unusual demand.
Learning rate	The weight of the previous hourly demand denoted in a percentage (e.g., 0.1%).
Accuracy replenishment moments	A binary variable: can be either yes or no. Indicates whether the replenishment moment can be determined with an accuracy of an hour on the day of issuing. See Section 3.2 for a detailed explanation.
Review moment	Either the result of continuous monitoring or specific moments in time on which an order can be placed (e.g., 11.00 AM). See Section 2 for a further explanation.
Order lead-time	The time between ordering and the actual cash upload. See Section 3.2 for further explanation.
Delivery weekdays	A list of all weekdays on which replenishments may be performed.
Delivery hours	Period of time during the day at which replenishments are performed
ATM capacity	Based on technical characteristics and the denominations available the ATM has a physical limitation of the amount of cash (e.g., € 260,000).
Interest rate	Determines the holding cost. The vendor does not receive interest yield from stored cash in the ATM. Costs are incurred indirectly because the cash inventory could yield profit if it was not stored.
Replenishment costs	The fixed costs of a single replenishment (e.g., € 120).
Target fill rate	The desired percentage of fulfilled withdrawals (e.g., 99%).

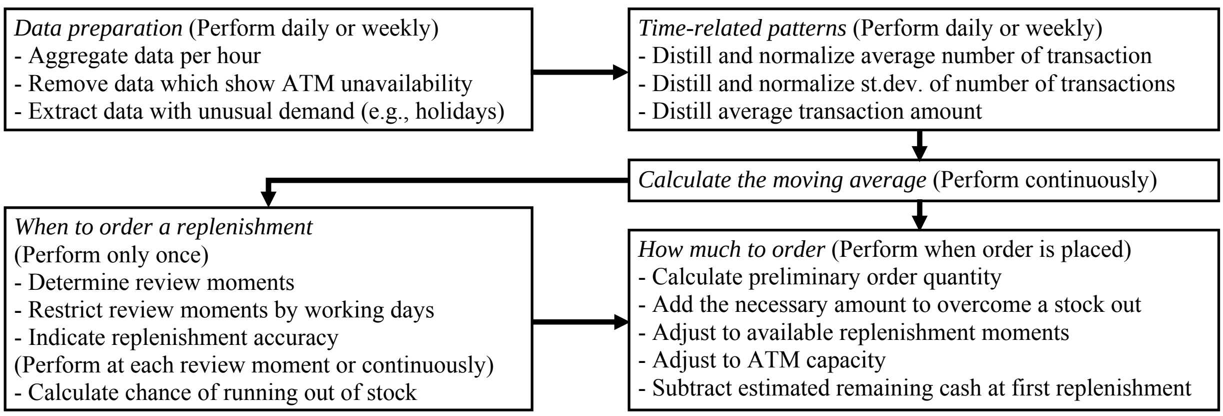


Figure 1. Schematic overview of the integrative model

Data preparation

In the first step, we show how to treat historical data such that time-based patterns can be identified for demand forecasting. We distinguish between the following steps:

A. Aggregate data by withdrawals per hour resulting in a dataset with the 'number of transactions per hour' and the 'average transaction quantity per hour'.
B. Exclude data points (i.e., records) from the dataset at which the ATM was unavailable based on an analysis of the dataset or by using log books depicting unavailability of the ATM.
C. Exclude data points that represent unusual demand in general such as the national holidays and surrounding days or regional events.

Step A is to aggregate all data per hour. Imagine a collection of data from all individual withdrawals over a period of two years, then the new dataset would consists of 2 * 365 * 24 = 17520 records after performing step A. Step B is performed by calculating a preliminary estimate of the demand per hour and examine whether there have been zero withdrawals while the estimate indicates that there should have been transactions. By checking the condition whether there have been zero transactions during a time period of 6 succeeding hours while the estimate indicates that more than twice the average hourly demand was expected, one can be relatively sure that the ATM was unavailable at the time period under consideration. Table 2 gives an example of ATM unavailability and shows which records need to be removed from the dataset. By trial and error we figured that 6 hours as minimum time period and 2 times the average hourly demand worked well.

Table 2. Example of excluding data points regarding ATM unavailability ("Y" = excluded)

	Dec. 28, 2008													Dec. 29, 2008										
Hour	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10
Real # of transactions	10	7	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	13	23
Est. # of transactions	11	14	17	18	16	15	18	15	10	8	5	4	2	1	0	0	0	0	1	2	3	7	13	16
Exclude from dataset?	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N

After removing all polluting data points regarding ATM unavailability, we continue in step C by identifying the 'unusual demand' resulting from the holidays in the area under consideration. Keeping track of regional events in time and location might be a lot of work. Clearly, the forecast will profit from excluding all data points from which it is known that these reflect unusual demand but in terms of time efficiency one might choose to exclude holidays and some surrounding days only. The excluded data should be used to forecast demand for these days in the future, which is described in the following step.

Filtering out time-related patterns

Continuing with the obtained dataset from the previous step, the next step of decomposition is to filter out time-related patterns to acquire an estimate arrival rate per hour and the average transaction amount per hour. The steps listed below are required. Notice that all four steps start with the dataset obtained from data preparation.

A. Distill and normalize the average number of transactions per weekday-hour, per month-day and per month
B. Distill and normalize the standard deviation of the number of transactions per weekday-hour
C. Distill the average transaction quantity per weekday-hour
D. Normalize the average number of transactions from days of unusual demand (e.g., holidays)

In step A we distill or filter out several time-related patterns sequentially. By calculating and normalizing the average number of transactions per weekday-hour (i.e., resulting in 24 * 7 = 168 values), we get insight in the weekday-hour pattern. Multiplying the obtained normalized values with the respective hourly demand figures in the dataset results in a new dataset in which the weekday-hour pattern no longer exists. The same procedure is repeated for the month-day pattern (i.e., 31 values) by making use of the newly obtained dataset without weekday-hour patterns. Next, the dataset without the weekday-hour pattern and without month-day pattern is used to filter out the month pattern (i.e., 12 values). The particular sequence – starting with weekday-hour, thereafter month-day and ending with month – is chosen by purpose to avoid problems with co-variance. For example, if the ATM is heaped on Saturdays and Sundays and these weekdays coincide with similar month-days for a couple of months, than the weekday pattern would be visible in the month-day pattern. This and similar problems are avoided to a large extent by starting with the pattern that explains most variance; the weekday-hour pattern.

Step B is to filter out a weekday-hour pattern from the standard deviations of the number of transactions per hour. Notice that we start with the dataset resulting from the previous step data preparation: In contrast to the calculation in Step A, data analyses at Dutch banks pointed out that the standard deviations do not follow a month-day or a month pattern. We calculate the average transaction quantity per weekday-hour in Step C according to the same procedure as in Step A and B. Similar to Step B, the average transaction quantity only follows a weekday-hour pattern and not a month-day or month pattern. In the first step (see the 'data preparation' step) we excluded days with unusual demand such as holidays. By normalizing these demand figures we can estimate demand for these days. We described how the time-related patterns should be filtered out of the dataset. In Figure 2 (see Section 4) the patterns are depicted using real life data of a single ATM. In step 3, the last part of the forecasting approach decomposition is described, which includes trend-information in the demand forecast.

Trend-information in demand forecasting

The demand forecast is required to adapt seamlessly to the continuously changing business environment. Data analyses at Dutch Banks showed that the demand over time follows a specific trend and that patterns (e.g., weekday-hour) are rather stable over time. Therefore, we include a moving average which is being updated hourly. After each hour, the deviation between the real and estimated number of transactions in the previous hour is calculated. This deviation – which can be positive or negative – multiplied with the learning rate and added to the moving average of the previous hour, results in a new average number of transactions per hour. The learning rate should be chosen somewhere in between 1% and 0.01%, depending on the stability of the demand.

To finalize this section we summarize the steps to calculate the estimated customer arrival rate and average transaction quantity:

- Customer arrival rate per hour*:* The moving average of the number of transactions per hour * the normalized values of the respective time-based patterns * the respective normalized holiday/event value. Notice that the approach also estimates the standard deviation of the customer arrival rate per hour.
- Average transaction amount*:* The average of all transaction amounts (in the prepared dataset) which fit a specific hour of a specific weekday.

Multiplying the customer arrival rate per hour with the average transaction quantity results in the estimated cash demand per hour.

When to order a replenishment

This step elaborates on the trade-off: When should an order for a replenishment be made such that the delivery will be performed just in time before the ATM runs out of cash?. We suggest to perform the following steps sequentially:

A. Make sure orders may only be placed on the review moments
B. Restrict the days of ordering by evaluating the working days of the replenishing company
C. Indicate the accuracy of the replenishment moment
D. Calculate whether not ordering at t_0 will result in a stock out; if so, place an order

The first step consists of making sure that orders are placed only on the agreed review moment. Cash centers require the order to be placed before the review moment in order to guarantee the agreed lead time. Next to the review moments, we have to deal with the working days of the replenishing company in step B. Ordering is inefficient when the replenishing company is not working on the estimated day of replenishment. So we have to subtract the lead time from their working days to determine the days on which ordering makes sense.

In our model we consider stochastic lead times, which means the actual moment of delivery is uncertain. The model is capable of distinguishing between two situations. In both situations the replenishing company agrees on a specific lead-time in days between the order placement and the cash upload. However, in the first situation the replenishing company is not promising a specific moment of replenishment during the day. So when calculating the reorder level, one should take into account that the actual replenishment could be performed in the early morning up to the late evening. In the second situation, the replenishing company agrees on a delivery accuracy of one hour. In the latter case the replenishment moment can be altered in the early morning on the day of issuing.

Given a specific review moment, the next review moment must be determined in the decision making process. Namely, if the next review moment is not after one day, but after more days (e.g., a weekend), then it is necessary to check – step E – whether not ordering at t_0 would possibly result in a stock out. In other words, we are interested in whether the inventory is sufficient to satisfy all the demand until the next delivery moment if we decide not to order at t_0. If inventory would be insufficient, it is clear that an order needs to be placed at t_0. If inventory would be sufficient, ordering can be postponed until the following review moment. However, calculating whether inventory would be sufficient (i.e., larger than reorder level) is rather complex because of the stochasticity involved. To be relatively sure that inventory is sufficient, the standard deviation of the demand is used. The chance of a stock out decreases by increasing the reorder level with a number of standard deviations. The model calculates a level of confidence, i.e., the required number of standard deviations, to make sure that the target fill rate is achieved. So, changing the input parameter 'target fill rate' allows for balancing between service and costs aspects.

How much to order

Once the ultimate decision is made to order a replenishment, the decision 'how much to order' needs to be made in succession. In this decision we try to minimize the long-run holding cost and the replenishment cost. There are two major drivers affecting the decision: The fixed cost per replenishment and the interest rate. We take the following steps:

A. Calculate preliminary order quantity by an iteration algorithm
B. Add up the necessary amount to overcome a stock out at the next replenishment
C. Restrict the order quantity to adjust to available replenishment moments
D. Restrict the order quantity to deal with the physical ATM capacity
E. Subtract the estimated remaining cash at the first replenishment from the order quantity

Note that we refer to the first replenishment – which is the replenishment that takes place when the lead time elapses – and the next replenishment which is the estimated succeeding replenishment after the first replenishment.

To find the replenishment quantity where long-run costs are minimal, we start in step A with calculating the average cost – for both replenishment and inventory holding – per hour for a replenishment quantity that equals the forecasted demand of one hour into the future. We continue calculating the average cost per hour by increasing the replenishment quantity with the hourly demand forecast for each succeeding calculation. When a minimum average cost is found, calculations are continued for 200 hours to make sure we did not get stuck in a local minimum. This approach results in a replenishment quantity that corresponds to a minimal cost per hour. Before this quantity can be set as a definite replenishment quantity we have to perform a number of checks and possibly some adjustments accordingly. As a matter of fact, we try to plan the next replenishment by choosing a specific order quantity.

In the decision when to place an order we included a level of confidence to overcome stock outs. This level of confidence will often result in some remaining cash in inventory on the moment of replenishment. So the first adjustment – step B – is to make sure we add up the estimated remaining cash in stock on the moment of the estimated next replenishment to the order quantity.

The next step C, is to deal with the available replenishment moments. The order quantity must not equal an amount of cash for which it is estimated that the next replenishment will be required during night or during days at which no deliveries can be performed. If we do not restrict the quantity, a chance exists that the next replenishment has to be performed earlier than necessary which would result in lots of remaining cash in inventory at the next replenishment. The input parameter 'delivery hours' states during which timespan replenishments are performed. In step D we limit the order quantity to make sure the order quantity plus the remaining cash in inventory do not exceed the ATM capacity at the first replenishment. When the ATM reaches its physical capacity, the replenishing company has to take the excess cash back to the cash center which is a costly exercise. The final step E is to subtract the expected remaining cash in inventory on the moment of the first replenishment.

4. COMPARING DECOMPOSITION AND ARTIFICIAL NEURAL NETWORKS

Simutis *et al.* (2008) report that in their opinion the best approach for ATM demand forecasting is an Artificial Neural Network approach (ANN). However, the authors did not compare it with decomposition. Consequently, we opt for first validating this component of our integrative model by comparing it with an ANN approach. Several important elements of the ANN approach of Simutis *et al.* (2008) are not reported for and therefore, we develop a new ANN approach to be used in several scenarios. We discuss decomposition and ANN sequentially and compare results afterwards.

Transaction data are used from a single ATM of a Dutch bank over a period of 1.5 years consisting of 143,468 registered withdrawals performed between January 1, 2008 and July 31, 2009. This ATM can be considered to be representative for most ATMs. The first step is to split up the dataset in a training set and a holdout set. The training set is used to filter out patterns and the holdout set serves the purpose of testing the quality of the approach. The first 15 months are added to the training set, the remaining 4 months to the holdout set. We prepared the dataset and filtered out the time-related patterns as is mentioned in the first two steps of the integrative model (see Section 3.2 for an explanation).

Figure 2 shows the weekday-hour, month-day and month patterns with the mean and standard deviation of the demand. Also the average transaction amount with standard deviation is shown. It is interesting to see that especially Wednesdays, Fridays and Saturdays show different patterns, that an increase in demand and standard deviation can be noticed during payday and that the demand during summer months is slightly larger.

The established ANN is a Multilayer Perceptron Network with the demand (in €) per hour as dependent variable and with 3 nominal factors: weekday-hour, month-day and month. The ANN architecture has one hidden layer with a varying amount (11 up to 19) of units. Training is done online with the gradient descent optimization algorithm with a momentum of 0.9, an interval center of 0, an interval offset of ± 0.5 and a maximum training time of 5 minutes. Other training characteristics concerning the learning rate are altered between scenarios and depicted in Table 3. The dataset is divided in three parts: 12 months for training, 3 months for testing and 4 months as holdout set.

The hourly forecasts are obtained over the last 4 months (corresponding to the holdout set) using both forecasting methods. The hourly forecasts are compared with the real demand during the same period and the mean absolute percentage error (MAPE) is used as performance measure. The MAPE of the obtained forecasts of both approaches are included in Table 3. The results show that decomposition outperforms the ANN. The difference between decomposition (37.79%) and the fifth ANN scenario (42.42%) is significant with a 99% confidence level. Although only one ATM and some ANN characteristics have been used we consider the results to be reliable.. We used a standard representative ATM and included a substantial amount of data (121,028 transactions) in the training set. Choosing different ANN settings can obviously change the performance of the ANN, but in our research we did not come across settings that could perform better than the fifth scenario.

We demonstrated that the forecasting approach decomposition (i.e., the first three steps) of the integrative model, provides more accurate estimates of future demand than the proposed ANN. In Section 5 we test the performance of the whole integrative model including decomposition in terms of robustness.

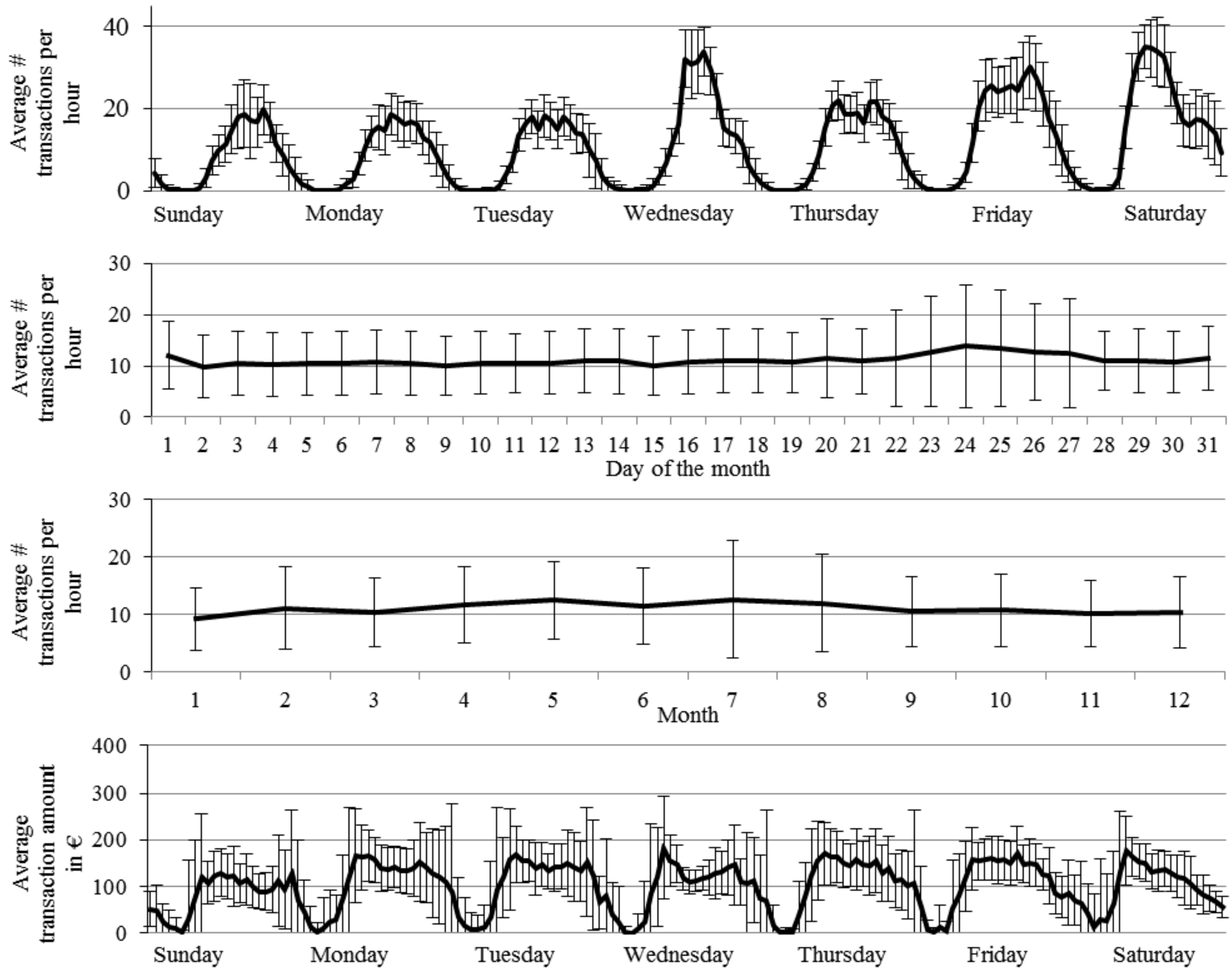


Figure 2. Mean ± the standard deviation of number of transactions per hour in weekday-hour (1st), month-day (2nd) and month (3rd) patterns, and the transaction amounts (4th)

Table 3. Results of the comparison between Decomposition and ANN

	Decomposition	ANN #1	ANN #2	ANN #3	ANN #4	ANN #5	ANN #6
Initial Learning Rate (LR)	N/A	0.001	0.001	0.005	0.001	0.01	0.05
Lower Bound LR	N/A	0.0001	0.00001	0.0005	0.0001	0.001	0.005
LR reduction, in epochs	N/A	10	10	10	50	10	10
MAPE (in %)	37.79%	42.42%	49.02%	44.63%	42.04%	40.51%	41.06%

5. ROBUSTNESS OF THE INTEGRATIVE MODEL

The second step in the validation stage is to test the robustness of the integrative model. The model has been implemented in Rockwell Arena 11 Simulation software. We inserted input parameters provided by commercial Dutch banks in the base model to reflect reality. Transaction data are obtained from the same ATM as introduced in Section 4. We test the robustness by comparing the base model with 17 scenarios for which varying values of the input parameters are chosen. The model can be considered robust when good results are also obtained when extreme values of the input parameters are chosen. coincide explicable performance indicators.

Table 4. Results of simulation experiments with the integrative model (over period of 4 months).

		Input parameters		Performance indicators				
		Base model	Scenarios	Total cost	Average cash inventory	Average # of replenish-ments	Actual fill rate (%)	Total amount excess cash due to ATM capacity
Base model	**0**	**N/A**	**N/A**	**€ 3,956**	**€ 118,631**	**16.4**	**99.78%**	**€ 3,478**
Demand	1	9/hr	12/hr	€ 4,528```	€ 129,337```	19.7```	99.68%```	€ 23,163```
(avg. cust./hr)	2		6/hr	€ 3,226```	€ 109,875```	11.6```	99.69%``	€ 702```
Sigma Demand	3	σ	2 * σ	€ 3,955	€ 118,517	16.5	99.77%	€ 2,838
	4		σ / 2	€ 3,958	€ 118,735	16.4	99.82%	€ 3,756
Trend in demand	5	N/A	+0.1%/day	€ 4,129```	€ 122,283```	17.4```	98.81%```	€ 5,674```
	6		-0.1%/day	€ 3,592```	€ 119,163	13.3```	99.83%`	€ 4,076
ATM Capacity	7	€ 260,000	€ 180,000	€ 4,130```	€ 92,647```	21.5```	99.67%```	€ 9,478```
	8		€ 400,000	€ 3,988```	€ 119,480	16.6```	99.78%	€ 0```
Interest rate	9	5%	1%	€ 2,163```	€ 133,272```	14.3```	99.76%	€ 12,208```
	10		10%	€ 5,671```	€ 105,799```	17.8```	99.54%```	€ 77```
Delivery days	11	Not at Sun	Not at Sat	€ 3,973```	€ 118,969	16.5``	99.78%	€ 3,958
	12		Also Sun	€ 3,955	€ 118,634	16.4	99.78%	€ 3,347
Replenishment cost	13	€ 120	€180	€ 2,835```	€ 105,799```	17.8```	99.54%```	€ 77```
	14		€ 60	€ 4,871```	€ 127,907```	15.2```	99.76%	€ 13,645```
Target Fill Rate	15	99.8%	99.6%	€ 3,931```	€ 117,882``	16.3``	99.60%```	€ 4,491
	16		99.9%	€ 3,976```	€ 119,087	16.5```	99.88%```	€ 3,261
Accurate repl. mom.	17	No	Yes	€ 3,851```	€ 112,238``	16.5	99.98%```	€ 598```

To generate customers and customer withdrawals in the simulation, we draw values from distributions obtained in Section 4. For ease of analyzing, we consider one review moment per day in the experiments. Actual replenishments are being performed between 8.00 AM and 5.00 PM with a uniform distribution. When more accurate delivery is available (can be indicated by the input parameter 'Accurate replenishment moment'), then replenishment is performed within a specified hour. ATM unavailability due to technical ATM failures is not taken into account, this means customer withdrawals can only be rejected when the ATM is out of stock. The target fill rate is set to 99.8%.

Based on the formulas presented in Law and Kelton (2000) we have calculated that a replication size of 100 is sufficient for all experiments in this paper. The results are presented in Table 4. The input parameters of the base model are listed in the first column and its performance is shown in the first row. Furthermore, the results of the other 17 scenarios are demonstrated in the succeeding rows. The second column displays how the respective input parameter is altered for each scenario. The performance of each scenario is compared with the base model and the quotation marks indicate the level of significance: ` = 90%, ``=95%, ```=99%. No quotation mark indicates the mean is indifferent from the base model. In the base model a total amount of € 3.3 million is requested by 26,746 withdrawals.

The base model shows a total cost of € 3,956 with an average inventory level of € 118,631, a fill-rate of 99.78% and a total amount of € 3,478 excess cash over a period of 4 months. The target fill rate is almost equal to the actual fill rate and the average inventory is fairly between zero and the ATM capacity. A good thing is that the amount of excess cash at replenishments (see last column in Table 4) is small in all scenarios, especially when considering the total amount of cash requested (€ 3.3 million). Most scenarios show expected and desired results, which are scenarios 1-4, and 6-17. For example, when the ATM capacity is smaller (scenario 7), more replenishments are required and the average inventory level is smaller. A somewhat surprising result is that a larger ATM capacity does not influence the average inventory level and the number of replenishments much. Apparently the presumed interest rate of 5% prevents a higher average inventory level. It can also be noticed that if placing a replenishment is, or is not possible during the whole weekend (see respectively scenarios 11 & 12) the performance indicators show almost no deviation. Apparently the model is capable of taking care of this restriction very well by choosing the right ordering quantities by which replenishments are not required to be performed on non-working days. Scenario 17 also provides insightful results: It shows that being able to determine the moment of replenishment on the day of issuing with an accuracy of an hour increases performance greatly: With a fill rate of 99.98% and a significant decrease in cost the last scenario easily outperforms the base scenario.

When a positive trend in demand is included (see scenario 5), we see unexpected results. All indicators differ normally, except for the actual fill rate. The model should adapt to the increasing demand, such that the target fill rate equals the actual fill rate. However the actual fill rate is a percentage lower than the target.

In general, most scenarios show desired results which indicate that the integrative model performs well in a variety of circumstances. We also point out that additional analysis is required to make sure that the target fill rate is always achieved, also when a positive trend in demand is presumed.

6. CONCLUSIONS AND FURTHER RESEARCH

This paper presents a novel dynamic solution approach that considers both ATM demand forecasting and ATM replenishment in an integrative way. We argue that integration of both aspects enhances the cash management model. In our integrative model, we consider continuous recalculation of the reorder level, based on real time demand forecasting. This means a replenishment is triggered by simultaneously evaluating the reorder level and the inventory level. Therefore, the integrative model is capable to deal with either periodic or continuous reviewing. Next to that the model considers lost-sales, variable order quantities, fixed replenishment cost, stochastic lead times and a target fill rate.

The model is capable of processing transaction data which allows for accurate and reliable forecasting. Decomposition is introduced as an important element in the forecasting component of the new integrative approach. Decomposition can be classified as a (greatly) modified linear regression approach with periodic coefficients. Computational results demonstrate that decomposition outperforms an artificial neural network approach which is considered to be a promising new technique to forecast time-series demand figures in ATM management. A sensitivity analysis shows the whole model performs well in many settings which shows general applicability of the model. We demonstrate that specific drawbacks of linear regression models do not apply to our method whereas the control parameters can be updated continuously and the approach is easily replicated.

Further research should include more exhaustive numerical experiments to check the practical applicability of the model proposed. Next to that, the model could be extended to consider rounded order quantities to correspond to an order of real banknotes. Another extension might be to develop a model in which inventory is managed for all individual denominations.

7. REFERENCES

Al-Saba, T. & El-Amin, I. (1999). Artificial neural networks as applied long-term demand forecasting. *Artificial Intelligence in Engineering, 13:* 189-197.

Armstrong, J. S. & Green, K. C. (2005). Demand forecasting: Evidence-based Methods. *Working paper Monash University*.

Bijvank, M., Bhulai, S. & Huh, W. T. (2010). Replenishment policies for lost-sales inventory systems with periodic reviews and order cost. *Working paper submitted for publication*.

Bijvank, M. & Vis, I. F. A. (2010). Lost-sales inventory theory: a review. *Working Paper VU Univeristy Amsterdam*.

Castro, J. (2009). A stochastic programming approach to cash management in banking. *European Journal of Operational Research, 192:* 963-974.

Harris, F. (1913). How many parts to make at once. *The Magazine of Management, 10*.

Hill, R. M. & Johansen, S. G. (2006). Optimal and near-optimal policies for lost sales inventory mdels with at most one replenishment outstanding. *European Journal of Operational Research, 169:* 111-132.

Johansen, S. G. & Hill, R. M. (2000). The (r,Q) control of a periodic-review inventory system with continuous demand and lost sales. *International Journal of Production Economics, 68:* 279-286.

Law, A. M. & Kelton, W. D. (2000). *Simulation Modeling and Analysis*, McGraw-Hill, New York.

Simutis, R., Dilijonas, D. & Bastina, L. (2008). Cash demand forecasting for ATM usting neural networks and support vector regression algorithms. *20th EURO Mini Conference,* 416-421.

Simutis, R., Dilijonas, D., Bastina, L., Friman, J., Drobinov, P. (2007). Optimization of cash management for ATM network. *Information Technology and Control, 36:* 117 - 121.

Tijms, H. C. & Groenevelt, H. (1984). Simple approximations for the reorder point in periodic and continuous review (s,S) inventory systems with service level constraints. *European Journal of Operational Research, 17:* 175-190.

TOWARDS ADVANCED LEARNING IN DISPATCHING RULE-BASED SCHEDULING

Bernd Scholz-Reiter, Jens Heger*, Torsten Hildebrandt and Daniel Rippel

BIBA - Bremer Institut für Produktion und Logistik GmbH
at the University of Bremen
Hochschulring 20
28359 Bremen, Germany
e-mail: {bsr|heg|hil|rip}@biba.uni-bremen.de

Abstract: Decentralized scheduling with dispatching rules is applied in many fields of logistics and production, especially in semiconductor manufacturing, which is characterized by high complexity and dynamics. Many dispatching rules have been found, which perform well on different scenarios, however no rule has been found, which outperforms other rules across various objectives. To tackle this drawback, approaches have been proposed, which select dispatching rules depending on the current system conditions. Most of these use learning techniques to switch between rules regarding the current system status. Studies of, e.g. Rasmussen (1996), show that Gaussian Processes as a machine learning technique can outperform other techniques like neural networks. We therefore investigate their use in the context of scheduling to select dispatching rules in dynamic scenarios. Our analysis has shown that Gaussian processes predict dispatching rule performances very well, making them well-suited as a learning technique in this application. Additionally, the advantage of Gaussian processes to provide a measure of prediction quality can be used to build a learning model incrementally. This can be used to save simulation effort.

1. INTRODUCTION

In today's highly competitive, globalized markets, an efficient use of production resources is inevitable for manufacturing enterprises. Therefore, especially capital-intensive industries like semi-conductor manufacturing spend considerable effort to optimize their production processes and, as one part of it, optimize production scheduling. Scheduling, as many other problems in the manufacturing domain are combinatorial, NP-hard optimization problems. It has attracted researchers for many decades now and is still of big interest, because of its high relevance and difficulty. Increase in computational power and continuous improvements in optimization and machine learning methods lead to more and more sophisticated heuristics.

One class of scheduling heuristics are dispatching rules which are widely used in industry, especially in complex manufacturing systems like semiconductor manufacturing. Their popularity derives from the fact that they perform reasonably well in a wide range of environments, and are relatively easy to understand. They also need only minimal computational time, which allows them to be used even in real-time, on-line scheduling environments taking into account the latest information available from the shop-floor. Dispatching rules as a special kind of priority rules are applied to assign a job to a machine. The dispatching rule assigns a priority to each waiting job each time a machine gets idle. This priority can be based on attributes of the job, the machines or the system. The job with the highest priority is chosen to be processed next. Dispatching rules have been developed and analyzed in the scientific literature for many years; see e.g. (Blackstone et al., 1982), (Haupt, 1989) and (Panwalkar et al., 1977). The most well-known rules are *Shortest Processing Time first* (SPT), *Earliest Due Date* (EDD) and *First In (Buffer) First Out* (FIFO).

Since the development of dispatching rules is a tedious and time-consuming, usually manually performed task, also concepts to generate them automatically have been proposed, e.g., (Hildebrandt et al., 2010). One drawback for sophisticated rules as well as for automatically generated rules remains: depending on the manufacturing system and the various objectives (e.g. mean flow time or mean tardiness etc.) no single rule, which outperforms all others can be found (Rajendran et al., 1999), (Mouelhi-Chibani et al., 2010).

To tackle this drawback of dispatching rules there have been a few approaches, which switch between rules depending on the system's conditions. Usually it is not possible to perform steady-state simulations with all rules under all occurring conditions due to the high dynamics and prohibitive computational cost. Therefore, statistical methods and learning methods have been proposed. Simulation runs are performed preliminary to generate learning data. These, usually very few, training cases are used to acquire knowledge, which can be used to make real-time decision during the actual production process.

* Corresponding author

One promising recent machine learning method are Gaussian Processes (GP), which have been introduced to the machine learning community in 1996 by Williams et al. (1996). Analysis showed that they performed very well in comparison to other techniques (Rasmussen, 1996). In addition to their good performance, they have the advantage, that they, besides estimating values, also provide a measure of their prediction quality. They are – their mathematical background aside – also relatively easy to handle.

For this reasons, in this paper we analyze the use of Gaussian processes for switching dispatching rules in production scheduling. As a first step we therefore perform a simulation study on a job-shop scheduling scenario introduced by Rajendran et al. (1999). The simulation runs are performed for different dispatching rules and system conditions with changing parameters. Afterwards the gained data enables the Gaussian Processes to make predictions, in which situation which rule performs best. We analyze how many samples are needed for a good prediction and the potential improvement, which can be achieved with this approach.

This paper is organized as follows: in section 2 we give a brief review of previous work on dispatching rules, machine learning in scheduling and Gaussian Processes. In section 3 our chosen scenario and the used framework are described. Section 4 presents the results of our experiments. The paper concludes with a short summary and gives directions towards future research.

2. STATE OF THE ART

2.1 Machine Learning in Scheduling

2.1.1 Scheduling

Haupt (1989) gives a definition of the scheduling or sequencing problem as "the determination of the order in which a set of jobs (tasks) {i | i = 1, ..., n) is to be processed through a set of machines (processors, work stations) (k | k=1...m).". Due to the high complexity, optimal solutions can only be calculated for small problems sizes. Therefore, especially in extremely complex scenarios with high dynamics dispatching rules are used. Depending on the current system state and the overall objective, there is no best rule, which outperforms all others. To select the best rule, the simplest way would be to perform steady-state simulations and determine which rule works best for the chosen scenario. In dynamic scenarios with changing system parameters, real-time solutions are needed. Hence, the use of learning methods to switch online between dispatching rules have been proposed.

2.1.2 Machine Learning Regression

"The goal of machine learning is to program computers to use example data or experience to solve a given problem" (Alpaydin, 2004). Here, we are interested in a system that can predict the value of the objective function which otherwise would have to be obtained by a costly simulation. As we are interested in a good due date-adherence of jobs, we chose the objective of mean tardiness. Inputs are the system's attributes (for example the utilization etc.) affecting the tardiness. The output is the estimated tardiness for the dispatching rule chosen to make scheduling decisions in our manufacturing system. Let X denote the (vector of) system attributes and Y the tardiness. Surveying past production processes (or using simulations) we can collect training data and the machine learning program fits a function to this data to learn Y as a function of X.

2.1.3 Machine Learning and scheduling: related research

Kotsiantis (Kotsiantis, 2007) gives an overview of a few supervised machine learning techniques, like artificial neural networks, decision trees, Naïve Bayes, support vector machines etc. Priore et al. (2001). present a review of machine learning in dynamic scheduling of flexible manufacturing systems. Most approaches are based on neural networks and are described in the following.

A simulation-based approach was presented by Wu and Wysk (Wu and Wysk, 1989). They switch regularly between different dispatching rule on machines. They proposed a multi-pass scheduling algorithm, which starts a short-term simulation of alternative rules and selects the best candidate for the manufacturing system.

A neural network based controller, consisting of an adjustment module and the equipment level controllers, was proposed for scheduling and controlling a manufacturing cell by Sun et al. (1996). The adjustment module considers the user objectives and the current performance levels to determine the relative importance of performance measures. Based on these importance values and current machine status, the equipment level controller, implemented by a neural network, selects a proper dispatching rule and the jobs are processed accordingly. The training samples for each equipment level controller are calculated by a one-machine simulation and modified to reflect the impacts of different dispatching rules on the system performance.

El-Bouri et al. (2006) used a neural network to select dispatching rule in a job shop. They chose small scenarios with 5 machines and investigated 3 rules. To train the neural network they calculated optimal solutions for 10, 15 and 20 jobs. The neural net was used to select one rule for every machine. With this approach they were able to get better results than just using one of the rules on every machine. The drawback of this approach is, that it is limited to scenarios with only a few machines, otherwise no optimal solutions for learning could be generated. Additionally, no dynamic scenarios and dynamic switching is considered.

Mouelhi-Ehibani and Pierreval use a neural network to dynamically switch dispatching rules on every machine depending on the current system state (Mouelhi-Chibani et al., 2010). They have selected four system parameters (e.g. shop load) and 22 system state variables (e.g. average slack time of jobs in the first queue), which the neural network uses to decide which rule should be applied. They train the neural net with preliminary simulation runs. The scenario they selected consists of only two machines and the set of dispatching rules consists of SPT and EDD. They outperform the static use of rules, but not that clearly, which might be caused by the small scenario.

These are interesting approaches, but the results seem to be improvable. It is not clear if this is due to the selected scenario or the learning technique. There has been no study on Gaussian processes for selecting dispatching rules until now. Since they have shown good results compared to other machine learning techniques (Rasmussen, 1996), this seems to be a promising approach.

2.2 Gaussian Processes

2.2.1 Introduction

O'Hagan (1978) represents an early reference from the statistics community for the use of a Gaussian process as a prior over functions, an idea which was only introduced to the machine learning community by Williams et al. (1996).

As stated before we have a simulation model implicitly implementing a (noisy) mapping between a vector of state variable (in our case containing, e.g. utilization) and the objective function (mean tardiness) $y = f(x) + \varepsilon$. The learning consists of finding a good approximation f*(x) of f(x) to make predictions at new points x.

To learn such a model using Gaussian processes requires some learning data as well as a so-called covariance function. This covariance function, sometimes called kernel, specifies the covariance between pairs of random variables and influences the possible form of the function f* learned. For our study the squared exponential (SE) covariance function is selected, because it should be able to fit to our data well and is a common choice in applications of GP. It is depicted in equation (1):

$$\mathrm{cov}\left(f\left(x_p\right), f\left(x_q\right)\right) = k\left(x_p, x_q\right) = \exp\left(-\frac{1}{2}\left|x_p - x_q\right|\right) \quad (1)$$

The formula shown is actually a simplified version of the formula used, which is given in formula (2). They differ by so-called hyperparameters. These parameters of a covariance function can be used to fine-tune the GP-model, thus learning of a GP model requires to have some learning data, choosing an appropriate covariance function and choosing a good set of hyperparameters. For further information see (Rasmussen et al., 2006) chapters 2 and 4.

2.2.2 Application and example

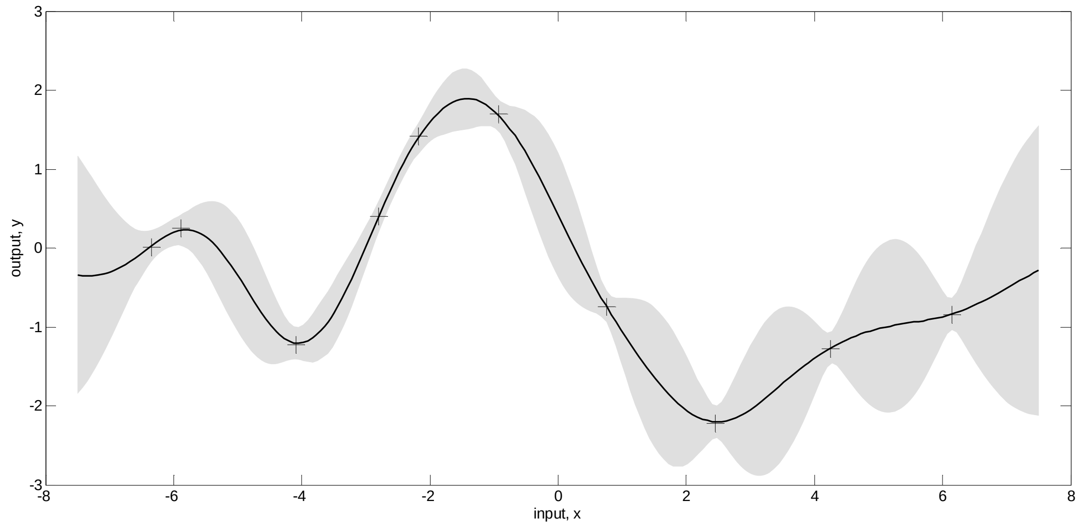


Figure 1. Example of a Gaussian Regression function with 10 noisy training points observed. The mean prediction is shown as a black line and the shaded area denotes twice the standard deviation

The *squared exponential* covariance function used in our experiments has three hyperparameters. There is the length-scale *l*, the signal variance σ_f^2 and the noise variance σ_n^2.

$$k_y\left(x_p, x_q\right) = \sigma_f^2 \exp\left(-\frac{1}{2l^2}\left(x_p - x_q\right)\right) + \sigma_n^2 \delta_{pq} \quad (2)$$

One big advantage of Gaussian processes compared to other machine learning techniques is that Gaussian processes also provide a quality estimate of their predicted value. This is denoted by the shaded area in Figure 1. Ten noisy training points are given and since there is noise the standard deviation close to the training points is small, but not exactly zero. In between two points as well as at the beginning and the end the quality of the estimates decreases.

Learning with Gaussian processes is done by selecting a covariance function and setting its free hyperparameters. For our study the *squared exponential* covariance function is chosen. To learn, or optimize the hyperparameters, the marginal likelihood should be maximized. Details and mathematical background can be found in (Rasmussen et al., 2006) chapter 5, especially equation (5.9) page 114). Basically, the hyperparameters are chosen in a way that the generalization error, which is the average error on unseen test examples, is minimized. This is done with cross-evaluation by splitting the training data in learning and test data. The training error is not optimized, because this may lead to over-fitting the data.

Additionally, since hyperparameters can be interpreted as length-scale parameters in the case of the *squared exponential* covariance function further optimizations can be performed. Rasmussen and Williams describe the hyperparameters informally like this: "how far do you need to move (along a particular axis) in input space for the function values to become uncorrelated". Thus, the *squared exponential* covariance function implements automatic relevance determination (ARD) (Neal, 1996), since the inverse of the length-scale determines how relevant an input is. A very large length-scale value means that the covariance will become almost independent of that input. ARD has been used successfully for removing irrelevant input by several authors, e.g. Williams et al. (1996).

3. EXPERIMENTAL SETUP

3.1 Problem description

The type of problems we address here, are dynamic shop scenarios with high complexity, like they are typical for semiconductor manufacturing. For our computational experiments we use the dynamic job-shop scenarios from (Rajendran er al., 1999). In total there are 10 machines on the shop floor, each job entering the system is assigned a random routing, i.e. machine visitation order is random with no machine being revisited. Processing times are drawn from a uniform discrete distribution ranging from 1 to 49. Job arrival is a Poisson process, i.e. inter-arrival times are exponentially distributed. The mean of this distribution is chosen to reach different desired utilization levels on all machines.

Following the procedure from (Rajendran et al., 1999) we start with an empty shop and simulate the system until we collected data from jobs numbering from 501 to 2500. The shop is further loaded with jobs, until the completion of these 2000 jobs to overcome the problem of censored data (Conway, 1965). Data on the first 500 jobs is disregarded to focus on the shop's steady state behavior.

The main focus of our research is to analyze the performance of Gaussian processes used for the selection of dispatching rules. This is interesting, because the major drawback of dispatching rules is that they lack a global view of the problem, i.e., they approach the overall scheduling problem by taking independent scheduling decisions based on the current, local conditions at the particular machine without consideration of the negative effects they might have on future decisions and on the overall objective function value.

3.2 Investigated dispatching rules

To have a set of dispatching rules out of which the best for each system's condition can be selected, we have selected some dispatching rules from the literature. The first five are being standard rules used for decades now, the sixth rule was developed by Rajendran and Holthaus especially for their scenarios and minimizes mean flowtime very well. If the rules calculate the same priority for more than one job, we use First-In-System-First-Out (FSFO) as a tiebreaker.

1. **SPT** – *Shortest Processing Time First*: SPT breaks ties by choosing the job with the shortest processing time for its imminent operation. Although this rule primarily aims to reduce the flow time of jobs (the difference between its completion and release time), SPT has shown to effectively minimize total tardiness when most jobs cannot meet their due dates, because of a tight due date settings and/or a high shop utilization.

2. **EDD** – *Earliest Due Date*: EDD resolves ties among equally weighted jobs by prioritizing the job with the earliest due date, therefore tending to decrease the maximum tardiness of all jobs. Contrary to SPT, the EDD rule is known to perform well for total tardiness when the shop is not congested and most jobs can be completed on-time.

3. **FSFO** – *First In System First Out*: Between jobs having identical weights, FSFO selects the job that has been the longest time in the system. Though FSFO generally exhibits a modest tardiness performance, it is easy to implement and often serves as a benchmark in simulation studies.

4. **FBFO** – *First In Buffer First Out*: FBFO is very similar to FSFO. The difference is that jobs are ordered after the arriving time in the system, but at the current buffer. Its performance is very equal to FSFO.

5. **MOD** – *Modified Operation Due Date*: MOD orders the queue of waiting jobs by the larger of each job's operation due date or each job's earliest possible finish time, whichever is larger. Therefore, if all jobs in the queue have positive slack (no job is in danger of missing its due date), then MOD dispatches them in earliest due-date (EDD) order. If all jobs have negative slack (all jobs are in danger of missing their due dates), then MOD works like SPT to reduce shop congestion. The weighted version of MOD is defined as follows:

$$WMOD_i = \frac{1}{W_i} \max(p_{i,imt}, d_{i,imt} - \tau) \tag{3}$$

6. **2PTPlusWINQPlusNPT** – *$2 \cdot$ Processing Time + Work In Next Queue + Next Processing Time*: This rule was suggested by Holthaus and Rajendran and consist of three parts (Holthaus et al., 2000). First, the processing time on the current machine is considered. Secondly, the *Work In Next Queue* is added:
   WINQ: jobs are ranked in the order of a (rather worst case) estimation of their waiting time before processing on the next machine can start. This estimation includes the time needed by a machine *m* to finish its current job plus the sum of processing times of all jobs currently waiting in front of *m*. The job where this sum is least has the highest priority. Thirdly, the processing time of a job's next operation (**NPT** – *Next Processing Time*) is added.

3.3 System architecture

For the simulation experiments we have used a Java-port of the job shop implementation of the SIMLIB library (Law, 2007), as described in (Huffman, 2001). This is a discrete-event simulation, which we have used to implement the manufacturing scenarios and assess the performance of the dispatching rules. Our framework allows the utilization of multi-core processors/computers, which was used in our experiments to run them on an 8-core computer with Intel Xeon 3GHz-CPUs.

For the Gaussian processes, we have used the software examples provided by Williams (2006) and adapted them for our scenarios. The calculations have been performed with MatLab from MathWorks.

3.4 Gaussian process regression for dispatching rule selection

The results of preliminary simulation runs are taken as learning data, which is used by the Gaussian processes to estimate the tardiness values of the dispatching rule. For every dispatching rule a Gaussian process regression is performed and the one with the lowest estimated mean tardiness is selected. To determine the estimation quality of the Gaussian processes we have split the 270 data points in three groups for cross-evaluation. Out of two groups (max 180 points), three sets each of size 10, 20, 30, 50, 100 and 1 set of all 180 – mostly non-overlapping data points – are taken each time for learning. The left 90 data points are taken for testing.

As described before the GP give an estimate for every testing point and every rule. The one with the smallest estimated tardiness is chosen. Afterwards we have calculated the estimation error by summing up the differences of the selected rule's value and the best rule's value for this parameter combination. Let us assume we have to test the parameter combination 0.83 utilization and due date factor 3. The GP gives us estimates of 150 minutes for EDD, 70 minutes for 2PTPlusWINQPlusNPT and 71 minutes for MOD. Thus, we choose 2PTPlusWINQPlusNPT. To assess the error for this single case, we calculate the difference of the true tardiness value of 2PTPlusWINQPlusNPT (77.6 minutes) and the best rule, which is MOD, with 65.3 minutes mean tardiness, which would be an error of 12.3 minutes. This shows, that if the predictions are wrong, the wrong rule is chosen, unless the tendency of the prediction is correct. If the prediction would have been higher than 71 for 2PTPlusWINQPlusNPT, MOD would have been chosen. The errors for all 90 test values are summed up.

However, during our experiments we have experienced a few problems with the automatic adaption of the hyperparameters. Sometimes the algorithm, which minimizes the marginal likelihood (see Rasmussen et al.), was stuck in a local optima. The found hyperparameters lead to a few estimates, which have fallen out of aligment. To tackle this problem, we have performed a two step process. First we have started the automatic calculation of the hyperparameters for all three rules for each parameter setting with standard start parameters (x_1 length-scale $\ell_1=1$, x_2 length-scale $\ell_2=1$, the signal variance $\sigma_f^2=1$ and the noise variance σ_n^2=log(sqrt(0.1)). Since all three dispatching rules have usually similar functions patterns, their hyperparameters are also similar. Thus, in a second step we have performed the hyperparameter calculation again, but have used the average values from the first step as start parameters; which fixed the problem.

4. EXPERIMENTS AND RESULTS

4.1 Simulation runs

We selected two system parameters, which are the input for the Gaussian processes. The first is the system's utilization and the second is the 'dueDateFactor', which defines the job's due date tightness (job's due date is set to x- times the job's total processing time + system release time). These two system parameters have been combined to 270 combinations. We have performed simulation runs with system utilizations from 70% till 99% and have combined these with due date factors starting from 2 till 10 (30*9=270). The six selected dispatching rules described in 3.2 have been evaluated on all these parameter combinations. Our performance criterion is mean tardiness, but the general approach is applicable to other objective functions as well. Each result for each combination of utilization, due date factor and dispatching rule is the average of 200 independent replications to get reliable estimates of the performance of our stochastic simulation.

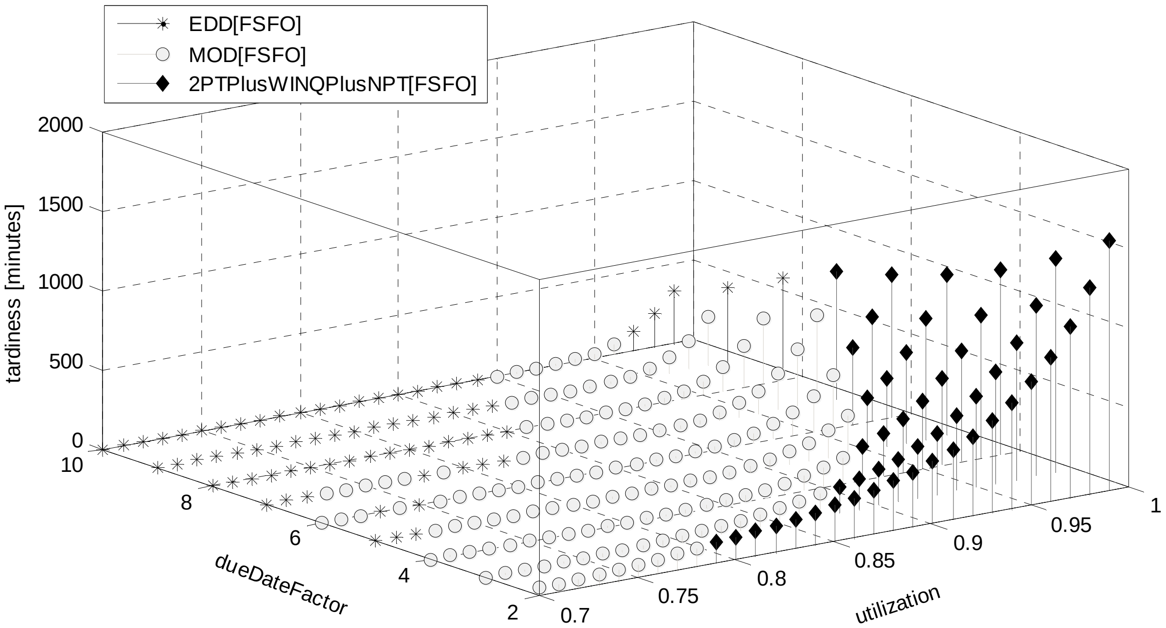


Figure 2. Results of simulation runs of 270 system parameter combinations. These reflect an optimal decision.

The results of these simulation runs are shown in Figure 2. Only best res ults for eac h combination are presented and the corresponding dispatching rule is marked. The rules SPT, FSFO and FBF O were always outperformed by one of the other three rules, so in the remainder of this paper only EDD, MOD and PTPlusWINQPlusNPT are used.

In scen arios wh ere t he due date is tig ht and th e utilization is high th e b est-performing ru le is 2PTPlusWINQPlusNPT. EDD is superior with relaxed due date settings and lower utilization rates. MOD performs well in m edium cases i n bet ween. T his sh ows t hat depe nding on t he sy stem's con ditions di fferent dispatching r ules outperform each other in the same scenario.

If simulation runs are expensive and more system parameters are considered not all parameter combinations can be simulated in ad vance. Th is i s wh ere sup ervised m achine l earning t echniques c an pl ay an i mportant r ole, helping to select the best dispatching rule with only a few simulations runs as a learning data set.

4.2 Benefits of rule switching

The interesting questions from the production scheduling point of view is to see, if learning with Gaussian processes can help to improve scheduling by selecting the best dispatching rule depending on the system's current conditions. An upper bound for this improvement is to first calculate the improvement possible with optimal switching (see Figure 2) compared with the strategy to always use the same fixed dispatching rules independently of the system's state.

The resuls are shown in Figure 4. As can be seen the best non-switching choice is to use th e MOD rule. Always using EDD is a rather bad cho ice resulting in an in crease of m ean tard iness by abo ut 800 0. Always using PTPlusWINQPlusNPT is almost as good as using MOD, resulting in an increase of about 1600 and 1500. This therefore sets the benchmark for the use of the learning method, as not to switch rules at all would lead to an increase of mean flowtime by 1500 if compared to the optimal switching decision.

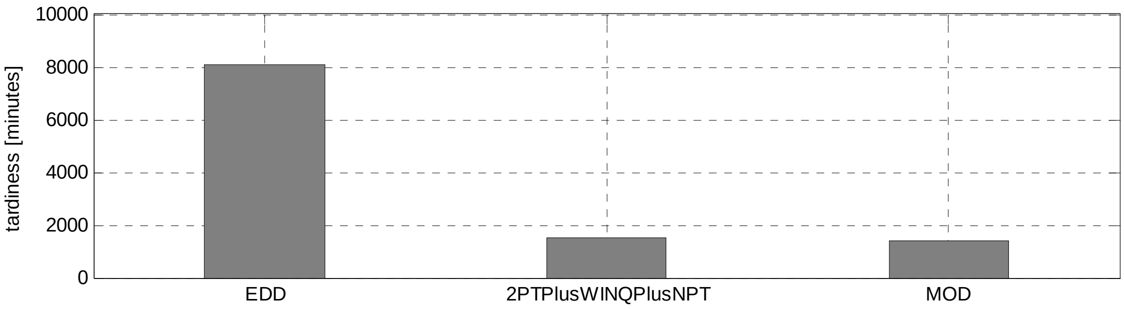


Figure 3. Difference in tardiness of only using a certain dispatching rule compared to always switching to the best rule

4.3 Estimation error and learning curves

The more data there is to learn for a Gaussian Process model the better the rule selection based on such a model should get. This intuition is confirmed by the graphs shown in Figure 3, showing the learning curves. These curves are generated using a 3-fold cross-validation of the data (for sample sizes smaller than 180: disregarding part of the learning data) and show the error compared with an optimal decision.

The results of Gaussian Process learning are very good. As can be seen there is a large improvement increasing the size of the learning data from 20 to 30. However, even 10 to 20 data points (i.e. simulations to generate them) are sufficient to improve over the best simple choice of always using MOD. With only 30 data points we were able to achieve 92 % of the improvements possible with switching dispatching rules based on the two attributes utilization and due date tightness. Using 180 data points yields about 99.5 %.

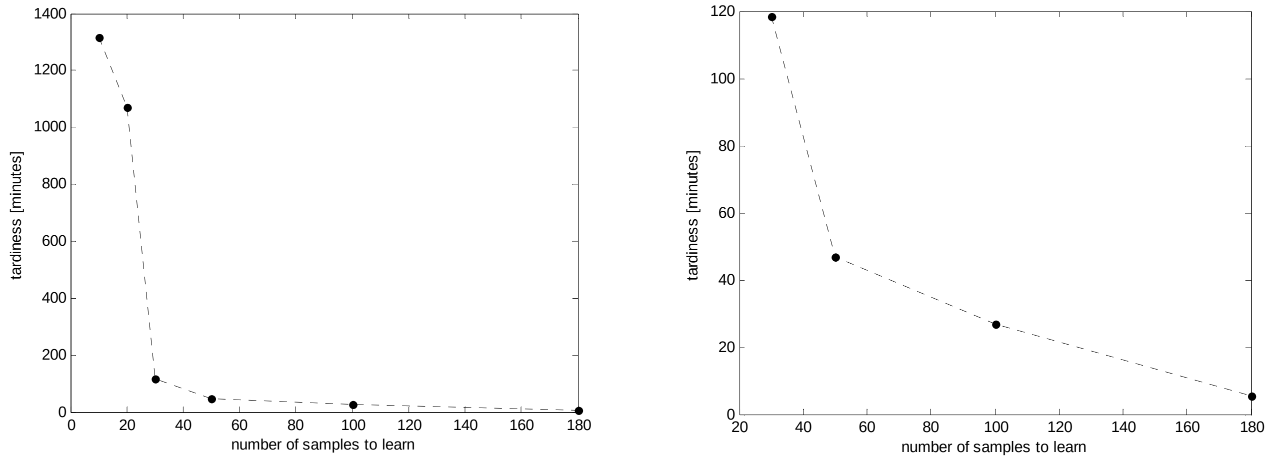


Figure 4. Learning curve – the overall error (sum of tardiness) in relation to the number of learning samples (left: all sample sizes (10 – 180); right: detailed view showing only sample sizes 30 – 180)

4.4 Dynamic Sampling and Continuous Model Improvement

The experiments presented in section 4.3 show the use of Gaussian Processes as a conventional supervised learning method: you have a certain amount of training data, then build a model of it and use this model later on. As shown before already 30 data points are sufficient to build a rather precise model. This number however largely depends on the problem and generating a learning curve as shown in Figure 4 might be too time-consuming if a single simulation run takes some time. However, using the uncertainty information we get from a GP model without extra effort, we can build a model incrementally, adaptive to the problem complexity and value ranges of the state descriptors actually needed. Therefore, we

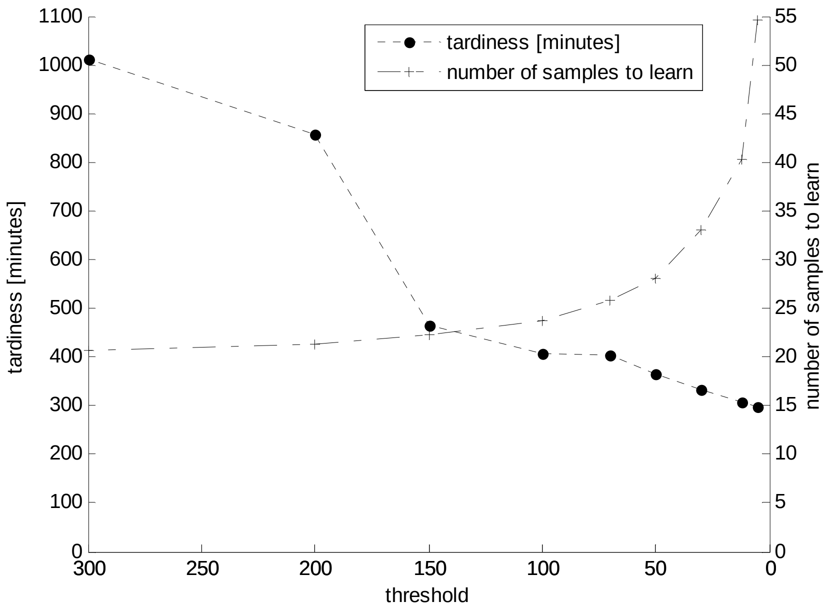


Figure 5. Learning curve of evaluation runs with dynamically increasing learning data

can use the simulation model and uncertainty information from the GP model to only simulate the data points required, i.e., to sample dynamically in contrast to building a static sample of learning data.

We start with a small sample size of 20 data points to learn the initial model. This model is then used to make predictions as before. However if the uncertainty associated with the performance of the best (according the models' predictions) rule at a certain point exceeds a certain threshold, we do a detailed examination (i.e. simulation) of this point and build a new, improved model from a learning data set amended by this point. This way we can continually improve the model until a sufficiently precise model is obtained.

In the following, we present experiments using this dynamic sampling strategy for different settings of the uncertainty threshold. As a measure of uncertainty, we use the standard deviation. The results of these experiments can be seen in Figure 5. The left axis and the associated tardiness graph show the solution quality associated with a certain threshold, whereas the right axis and sample size graph show the number of data points sampled. This is the number of simulations that would have to be performed, i.e. the initial learning data size of 20 plus the points dynamically sampled on all points where the uncertainty exceeds the threshold. As can be seen in the figure there are trends in opposite directions. Using a large threshold leads to a high tardiness but only very few points are sampled additionally. A small threshold causes a high number of data points to be sampled, but a high-quality model, i.e. low tardiness.

5. CONCLUSION AND OUTLOOK

In dynamic manufacturing scenarios with frequently changing system parameters adaptive scheduling approaches seem necessary. With this study we have shown that scheduling with dispatching rules can be further improved, if advanced learning techniques are used to adapt the scheduling mechanism to dynamic changes by selecting the best rule for the current system status. To our knowledge Gaussian processes have been used in scheduling for the first time. The estimation errors have been small for learning sample sizes bigger than 30, for our case with two system parameters. For more system parameters this approach should be scalable and make even more sense, because the number of simulation runs increases clearly then. If simulation runs are expensive and many parameter settings are possible not all cases can be simulated in advance. In these cases adaptive scheduling strategies with learning components seem promising. In such settings the approach described in section 4.4. to dynamically increase the learning data selectively with those points, where the prediction uncertainty is high, can also be used effectively. Providing such a measure of prediction uncertainty is a feature distinguishing GPs from most other machine learning techniques.

Our approach only conducts an analysis how GP could be implemented in scheduling. In future research our framework should be extended to a dynamic component, which selects dispatching rules in real-time. One advantage of Gaussian processes is that they give an estimate and its prediction quality. This could be used in the framework as well, for example to start new simulation runs for more training data, if the prediction quality is not satisfying.

ACKNOWLEDGEMENTS

The authors are grateful to the generous support by the German Research Foundation (DFG) under grant SCHO 540/17-1.

6. REFERENCES

Alpaydin, Ethem (2004). *Introduction to Machine Learning (Adaptive Computation and Machine Learning Series)*, volume 14. The MIT Press.

Blackstone, J. H., Phillips, D. T., and Hogg, G. L. (1982). A state-of-the-art survey of dispatching rules for manufacturing job shop operations. *International Journal of Production Research*, 20(1):27–45.

Conway, R. W. (1965). Priority dispatching and job lateness in a job shop. *Journal of Industrial Engineering*, 16:228–237.

El-Bouri, Ahmed, Shah, and Pramit (2006). A neural network for dispatching rule selection in a job shop. *The International Journal of Advanced Manufacturing Technology*, 31(3-4):342–349.

Haupt, R. (1989). A survey of priority rule-based scheduling. *OR Spektrum*, 11(1):3–16.

Hildebrandt, T., Heger, J., and Scholz-Reiter, B. (2010). Towards improved dispatching rules for complex shop floor scenarios - a genetic programming approach. In *Proceedings of the 12th annual conference on Genetic and evolutionary computation*, Portland, USA. (accepted paper, to appear).

Holthaus, O. and Rajendran, C. (2000). Efficient jobshop dispatching rules: further developments. *Production Planning & Control*, 11(2):171–178.

Huffman, B. J. (2001). An object-oriented version of SIMLIB (a simple simulation package). *INFORMS Transactions on Education*, 2(1):1–15.

Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica*, 31:249–268.

Law, A. M. (2007). *Simulation Modeling and Analysis*. McGraw-Hill, Boston, USA, 4th edition.

Mouelhi-Chibani, Wiem and Pierreval, Henri (2010). Training a neural network to select dispatching rules in real time. *Computers & Industrial Engineering*, 58(2):249 – 256. Scheduling in Healthcare and Industrial Systems.

Neal, Radford M. (1996). *Bayesian Learning for Neural Networks (Lecture Notes in Statistics)*. Springer, 1 edition.

O'Hagan, A. (1978). Curve fitting and optimal design. *Journal of the Royal Statistical Society*, 40(1):1–42.

Panwalkar, S. S. and Iskander, W. (1977). A survey of scheduling rules. *Operations Research*, 25(1):45–61.

Priore, Paolo, de la Fuente, David, Gomez, Alberto, and Puente, Javier (2001). A review of machine learning in dynamic scheduling of flexible manufacturing systems. *AI EDAM*, 15(03):251–263.

Rajendran, C. and Holthaus, O. (1999). A comparative study of dispatching rules in dynamic flowshops and jobshops. *European Journal of Operational Research*, 116(1):156–170.

Rasmussen, Carl E. and Williams, Christopher K. I. (2006). *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press.

Rasmussen, Carl Edward (1996). Evaluation of gaussian processes and other methods for non-linear regression.

Sun, Y. L. and Yih, Y. (1996). An intelligent controller for manufacturing cells. *International Journal of Production Research*, 34(8):2353–2373.

Williams, Chris (2006). Gaussian processes for machine learning - software examples. http://www.gaussianprocess.org/gpml/code/matlab/doc/.

Williams, Christopher K. I. and Rasmussen, Carl Edward (1996). Gaussian processes for regression. *Advances in Neural Information Processing Systems*, 8:514–520.

Wu, Szu-Yung David and Wysk, Richard A. (1989). An application of discrete-event simulation to on-line control and scheduling in flexible manufacturing. *International Journal of Production Research*, 27(9):1603–1623.

APPROACHES FOR REALIZATION OF AUTONOMOUS LOGISTICS IN PRACTICE

Jakub Piotrowski[1], Bernd Scholz-Reiter[2]

[1] Department Planning and Control of Production Systems,
University of Bremen,
c/o BIBA, Hochschulring 20,
Bremen, Germany,
email: pio@biba.uni-bremen.de

[2] Bremer Institut für Produktion und Logistik GmbH (BIBA),
University of Bremen,
Hochschulring 20,
Bremen, Germany,
email: bsr@biba.uni-bremen.de

Abstract: Today, planning and control of logistic processes is generally executed by centralized logistic systems, which cannot cope with the high requirements for flexible order processing due to increasing dynamics and complexity. In many cases this is caused by a non-synchronized information and material flow in logistic systems. There is an ongoing paradigm shift from centralized control of 'non-intelligent' items in hierarchical structures towards decentralized control of 'intelligent' items in heterarchical structures. Those intelligent items could either be raw materials, components or products (e.g. vehicles) as well as loading equipment (e.g. pallets, packages) or transportation systems (e.g. conveyors, trucks). The main characteristic of an intelligent item is its capability to control itself. This paper is dedicated to the idea of autonomous control as a new control paradigm for logistic processes. It focuses the results of the prototypical implementation of autonomous control mechanisms in automobile logistics.

1. INTRODUCTION

The term "autonomy" describes the capability of a system, process or object to design its input-, throughput- and output-profiles as an anticipative or reactive answer to changing environmental parameters. One specific criterion of autonomous processes or objects is to render a decision by itself on the basis of parameters (Scholz-Reiter et al., 2006). In the last years a paradigm shift in logistic processes has occurred. The centralized control of 'non-intelligent' logistic objects in hierarchical structures has been changed to decentralized control by 'intelligent' items in heterarchical structures. As shown in Figure 1 conventional logistic control systems are characterized by central planning and control processes, a hierarchical information structure. The hierarchical information structure does not allow fast and flexible adaptation to changing environmental influences, especially in logistic systems with a high degree of dynamic and complexity. In general dynamic in logistic systems can be described as the changes of a set of parameters and system variables in a defined time period. The behavior of a logistic system follows its intrinsic dynamics, which is influenced by the system parameters. As a result of the dynamics, the system variables change over time. One definition of dynamics could be: "the accelerated variation of the system status over time" (Scholz-Reiter et. al., 2002).

1.1 Autonomous control

Autonomous Control describes processes of decentralized decision-making in heterarchical structures. This requires that interacting elements in non-deterministic systems have the ability and opportunity for autonomous decision-making. The definition of autonomous control describes the maximum expression of the autonomous control in a logistics system. Consequently, all logistic objects in autonomous controlled logistic system would operate independently according to their own objectives. Autonomous control is given if a logistic object is able to process information, make decisions and execute the decisions by itself (Windt, 2005).

The aim of using autonomous control mechanisms is to increase robustness and positive emergence of logistics systems by a distributed and flexible coping with dynamics and complexity (Windt, 2007). In several simulation studies the use of autonomous control causes more flexibility, adaptivity and robustness to logistics systems. This is the case especially in logistic scenarios with a high degree of dynamic, for example caused by express orders or machine failures (Böse, 2005).

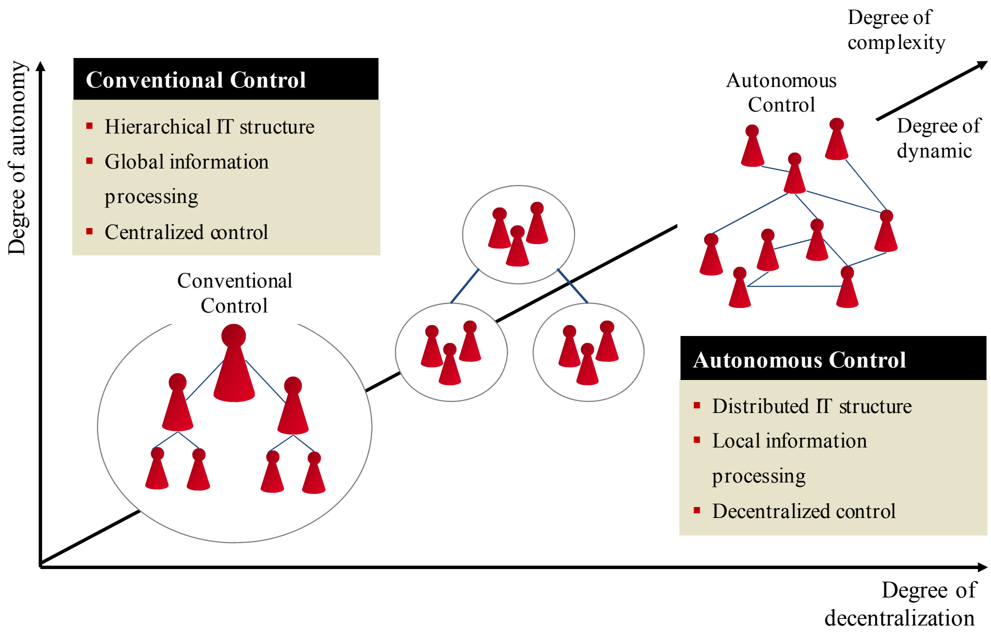


Figure 1. From conventional to autonomous control

In practice the degree of autonomy and the degree of decentralization will not have the highest value – in practice autonomous controlled logistic objects will be clustered in sub-systems. Each sub-system can operate autonomously and is characterized by the highest level of autonomy.

1.2 Information and communication technologies as an enabler for autonomous control

The dynamic development of information and communication technologies, e.g. the identification technologies like RFID (Radio Frequency Identification) or nearly everywhere available communication technologies like GPRS (General Packet Radio Service), makes intelligent processes (and therefore intelligent items or autonomy) possible. Furthermore the information and communication technologies become smaller and cheaper. Since January 2004 a German Collaborative Research Centre (CRC 637) funded by the German Research Foundation (DFG) has been established at the University of Bremen. It is named "Autonomous Cooperating Logistic Processes – A Paradigm Shift and its Limitations". Four faculties cooperate in this interdisciplinary project: engineering technology, economics, informatics, mathematics and electrical engineering. The research concentrates on modeling of autonomous cooperating logistic processes, designing methods and adequate tools as well as on evaluation for practical use.

In the field of logistics there are many remunerative uses of research results with information and communication technologies. There are prototypic stores where products, shelf, shopping carts and store cards are equipped with RFID chips (smart labels). Typically product RFID contains information about product type, durability and cost, the shelf about the amount and kind of its products, the shopping carts support the customer's navigation. The payment is completely automated because the checkout registers every product in the shopping cart (or the customer's pocket). Screens can show customer individual spots adjusted to the contents of the shopping cart. A shelf automatically orders replenishment when necessary. But there are many future possibilities.

The vision of the future could be decentralized distributed architectures of intelligent and communicating objects instead of today's centralized control of non-intelligent objects in hierarchical structures.: The implementation of these systems goes along with the use of existing and enhanced technologies. The flow of goods is no longer controlled by a central instance. Logistic objects like packages or vehicles can find their way through the logistic network to the destination autonomously (e.g. from the producer to the customer) while constantly communicating with conveyances and nodes and considering demands, e.g. concerning delivery date and costs.

2. REQUIREMENTS TO REALIZE THE IDEA OF AUTONOMOUS COOPERATING LOGISTIC OBJEKTS

The Idea of decentralized decision-making with autonomous control methods requires a real-time information flow to allow logistic objects to analyze the current system situation and use this data to decide, which next steps have to be done. To realize this idea, identification technologies like RFID, real-time locating systems like GPS and communication technologies like GSM or GPRS are needed. On the one hand, logistic objects are able to collect information about their own states. On the other hand, this information can be passed through to other logistic objects if required. In the real world there are many difficulties to realize the idea of autonomous control or autonomous cooperating logistic objects. In most cases logistic objects are not able to communicate, they have neither sensors that can be used to collect state data nor processing modules that can analyze this data and execute the decisions by themselves. Imagine a package that wants to route through a logistic system by itself. There is a couple of data required to realize the idea of autonomous control. There is also a need of information about available route infrastructure, available trucks at any transfer point or information about traffic situation on routes. Today the most logistic objects are not able to collect this data by themselves. In order to enable this, the use of additional information and communication technologies is necessary. There are many possibilities to expand non intelligent logistic objects to enable a self decision-making. In most cases the representation of these objects is realized be using multi agent system (MAS) (Böse, 2009). Therefore there is a need to identify a logistic object and represent it as an agent within a multi agent system. Secondary positioning data are needed to know where each object is within the logistics system. In case of a self routing package the multi agent system contains the information about the state of transfer points or available trucks. To collect the identification and positioning data several techniques can be applied. In several research activities RFID-Gates or mobile devices has been already used (Böse, 2009).

In case of using RFID-Gates to collect positioning data of logistic objects, there is a need to place the gates on each deciding place.

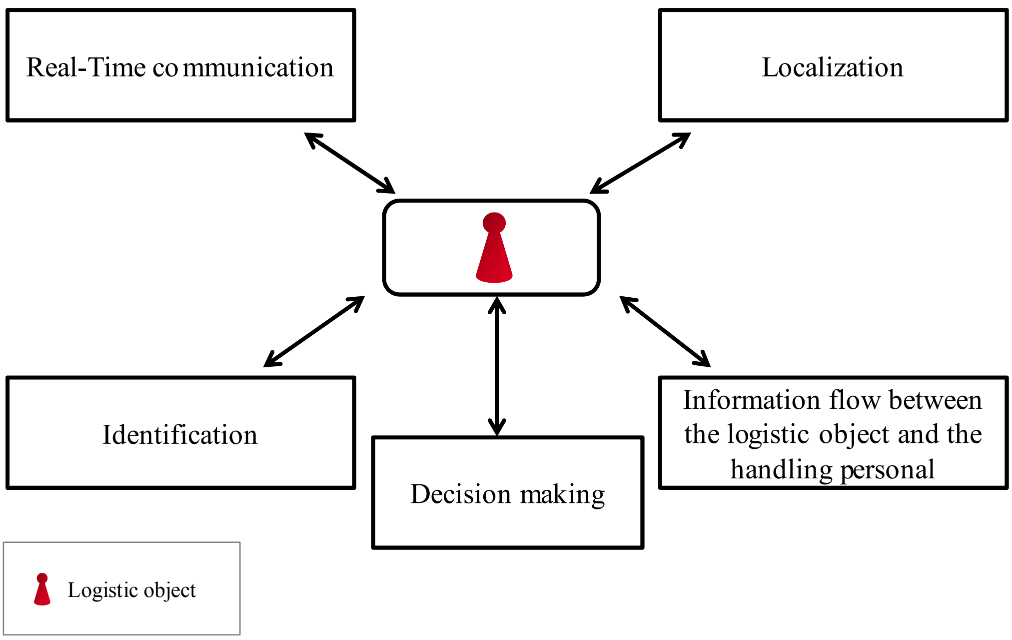


Figure 2. Components to realize autonomous control methods

2.1 Hardware related requirements

An appropriate technical hardware system has to fulfil different system requirements which are introduced in the following. Based on operational processes several system requirements for the hardware system can be deviated. These requirements – containing technical requirements, user requirements as well as safety requirements – have to be fulfilled to support staff in operational business and as a result to benefit from the introduced potentials for improvement.

Technical requirements

The Requirements relate to technological aspects of technical implementation with main focus on hardware and middleware. Summarized, the main technical requirements are:

- Operation time: Availability of a power management including standby function to reduce energy demand and consequently extend the operation time of rechargeable batteries.
- Robustness / shock resistance: High robustness and shock resistance combined with system reliability of hardware components as a protection against dropping down (up to a height of 1.5 m).
- Computing power: Providing adequate computing power of the mobile system (central processing unit, main memory) to guarantee short response times of the application software and as a result to avoid dead times waiting for completion of data processing.
- Cleaning: Easy and damage-free cleaning of applied, mainly outdoor hardware components (mobile data entry devices, reader gates etc.).

Safety requirements

The safety requirements contain several aspects including health and safety at work, data security as well as system reliability. The safety requirements can be described as follows:

- Electromagnetic radiation: Compliance with limit values regarding electromagnetic emission (EMA according to DIN 55022, DIN 61000-3-2 and DIN 61000-3-3) as well as electromagnetic interference (EMB according to DIN 55024).
- Data security: high data security regarding data exchange between mobile data entry device and logistic backend system (e.g. using encryption standard WPA (Wireless Protected Access)) in case of communication via WLAN), encoding of transponder data to prevent data manipulation or unauthorized readout of data, password protection of the mobile system to secure stored sensitive data (e.g. in the event of stealing or leaving the mobile data entry device inside a delivered car).
- User authentication: Unique user authentication for quality management purposes (allocation of users to executed activities, increase of data security).
- Operational reliability: High operational reliability (e.g. intermediate data storage to guard against power blackout, using emergency power supply).

User requirements

User requirements describe the requirements for the IT system from the user's point of view. These requirements concern particularly the handling of the application software as well as the mobile data entry device.

- Design of graphical user interface: Easy to use and quickly learnable handling of the application software, comprehensibility and clearness of terminal dialogs, permanent view of common information (e.g. strength of GPS and WLAN signal, status of battery charge, date and time as well as mobile data entry device number), definition of standard font size with respect to a viewing distance of 300 mm.
- Physical dimensions, weight and handling: Low weight (overall weight including battery < 500 g) and adequate physical dimensions of the mobile system to guarantee its manageability, operating temperature between 10°C and 50°C, dust and splash water protection in all weathers, one-hand operation.
- Keyboard: Use of a lighted keyboard to allow for good readability at night, definition of size and alignment of keyboard with respect to working conditions (usability with gloves etc.), key lock.
- Display: Image quality, reflection properties and colour scheme (in accordance to ISO 13406-2), good readability of dialogs presented on the display from an angle of view of 20° to 40° (in accordance to ISO 13406-2/7.2) and the keyboard labelling in all weathers (rain, snow, direct sunlight etc.), using an outdoor display, setting of contrast and brightness at a mobile system, backlight.
- User acceptance: Assistance of staff, low training time and effort, easy to use mobile system, no monitoring of employee activities, error-free operation of the mobile system, no limitation of employee competence caused by order processing automation, no impairment of health.

2.2 Software related requirements

Today the realization of the paradigm of autonomous control by giving the "intelligence" directly to the logistic object is often not reasonable form the economic point of view. The real-time representation using a mobile device makes economic sense only if the relevant logistic objects (e.g. packages) are passing through the logistic system frequently. As mentioned above it is possible to represent logistic objects in a virtual environment like a multi agent system to enable a decentralized decision-making of these objects. The representation could be done by the use of a mobile device. The operating staff could use mobile devices as a gateway or interface between the real-world and the virtual environment. In the face of the technical, user and safety requirements it is necessary to take care of the implementation design of the mobile device solution. A modular implementation of the required components and modules allows a flexible replace or addiction of hardware components among the development process. In general there is a need of components for identification, communication, positioning and a module for the control of the process data.

Figure 3 gives an overview of the relevant device components according to the mentioned requirements and needed hardware components for the realization of autonomous control methods. For each hardware cluster (like hardware modules for identification or positioning and sensors) a gateway is implemented within the hardware abstraction layer. The gateway has to realize the communication between the hardware module and the other hardware components. As shown in Figure 3 sensors like infrared sensors are clustered and handled by the sensor gateway. The processing unit has to organize the internal communication between the gateways. The logic layer coordinates the processes commands, makes logical decisions and evaluations as well as performs calculations. The main function of the power control layer is to turn off hardware components when they are not needed to realize an efficient power management. According to the required data for a autonomous decision making components are turned on. The necessary commands and functions to control the hardware modules are implemented in the logic layer. All data and information are stored and retrieved from an internal data base. The information and data are then passed back to the decision layer for processing and then eventually to the user by the utilization of the user interface.

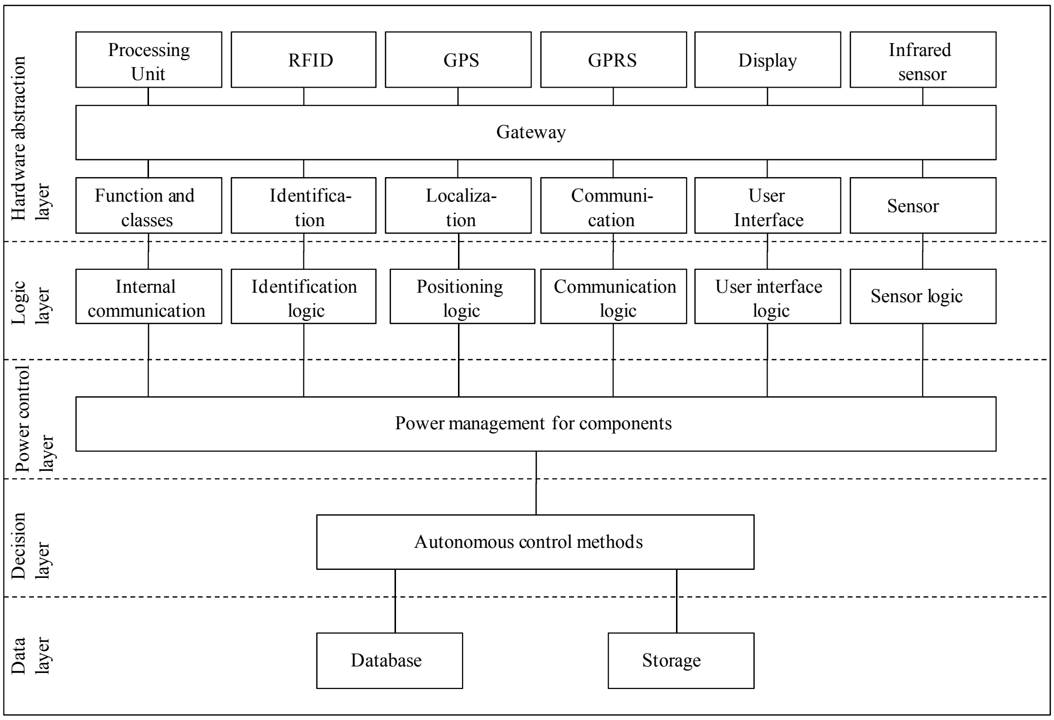


Figure 3. Software architecture

As shown in Figure 3 each system is represented in a multi agent system (MAS) by so called broker agent. The broker agent is responsible for the correct representation of each device which represents the real world object and the real world object data (including the object identification, specific object information, like type and color etc.). The MAS is coupled to the IT-backend system which contains all available data about for e.g. orders and delivery times. Specific data of the logistic object are stored within the MAS or the device for virtual representation.

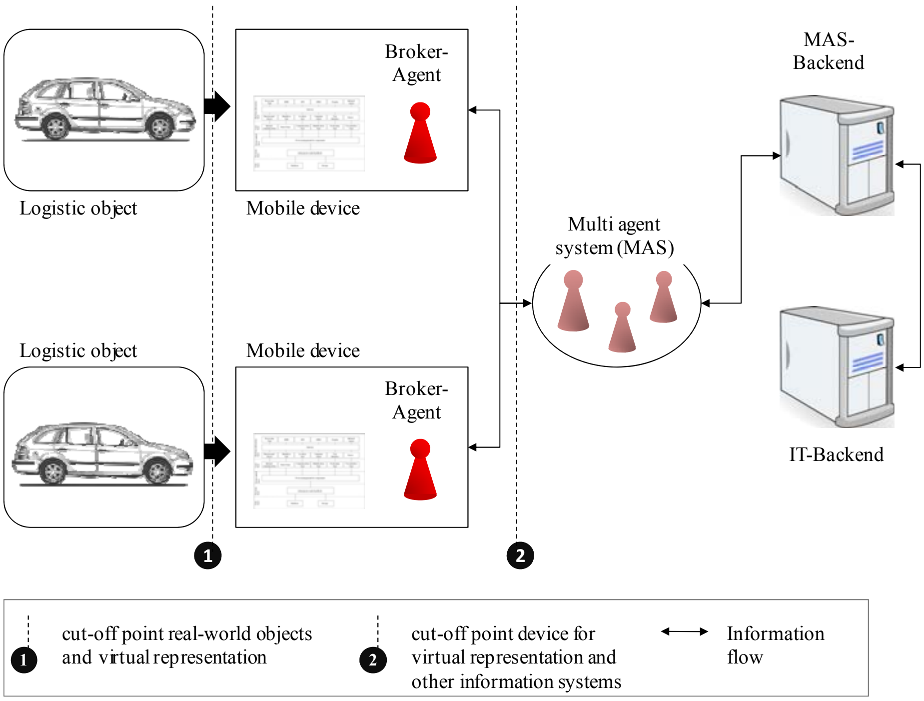


Figure 4. System concept

3. REALIZATION OF AUTONOMOUS COOPERATING LOGISTIC OBJECTS

As mentioned today there are several requirements that has to be fulfilled to realize the idea of autonomous cooperating logistic objects. The main challenge is to synchronize real-world data with the data within the MAS to allow logistic objects to make decentralized decisions by the use of the MAS. Therefore a software and system concept for a mobile solution was introduced above. This concept can be used on and with mobile devices for data entry in case of movable logistic object or on an additional unit on stationary machines or workplaces. Within the CRC 637 the technical development focused on the scenario of an automobile terminal and autonomous cooperating vehicles. The BLG LOGISTICS GROUP AG & Co. KG (BLG) develops and provides complex services for new and used vehicles in the range of transport, handling, technical treatment and storage. The BLG has established a Europe-wide logistics network on the basis of automobile terminals at strategically important traffic junctions. Every vehicle passes a set of process steps in the automobile logistics network: collection of vehicles at automobile manufacturer, multi-modal transport to automobile terminal via road, rail or inland waterway/sea, storage and technical treatment as well as delivery to automobile dealer. After delivery, each vehicle is identified by its vehicle identification number (VIN) from the terminal staff using mobile data entry devices which can read barcodes placed inside the vehicle behind the wind-screen. The VIN allows the assignment of the vehicle to its storage and technical treatment orders stored in the logistic IT-system. Based on predefined priorities, the IT-system allocates a storage location of a storage area to each vehicle. A handling employee moves the vehicle to the assigned storage location. After removal from stock, the vehicles possibly run through several technical treatment stations as fuel station or car wash. The sequence of the technical treatment stations is specified in the technical treatment order of the vehicle. Upon completion of all technical treatment tasks, the vehicle is brought to the shipment area for transportation to the automobile dealer (Böse et al., 2006 & Scholz-Reiter et al., 2008). The idea of autonomous cooperating vehicles focuses on the shift of the decision making from a central planning and control system to the vehicle. According to the idea of autonomous control, autonomous logistic objects are enabled to process information and render and execute decisions on their own. The use of autonomous control methods on automobile terminal show in simulations many advantages which were presented by (Böse et al., 2009 & Windt et al., 2009). In consequence, the vehicles and other logistic objects like storage areas have their own master data and act independently regarding their local objective system.

3.1 Wearable computing systems for realization of autonomous cooperating vehicles

The idea for further development contains the integration of different ICT components for identification, localization, communication and user interaction tasks in the work clothes of employees as a wearable computing system. The integration of electronic systems in clothing has been researched by many task forces (Kirisci et al. 2006). A special challenge is the high integration level of technologies as shown in Figure 5. Many of the previously developed systems were rejected by the users because of deficient ergonomics of the developed work clothes as well as manifold doubts about data security or impairment of health (high radiation of electronic components etc.). In the context of the mentioned scenario wearable computing systems have to be fixed closely to the body to avoid damages to the vehicles, e.g. damages of vehicles, while walking through storage places (Mrugala et al. 2009). Looking at this scenario, the idea of wearable computing seems to have more potential as common mobile solutions for the use while handling vehicles.

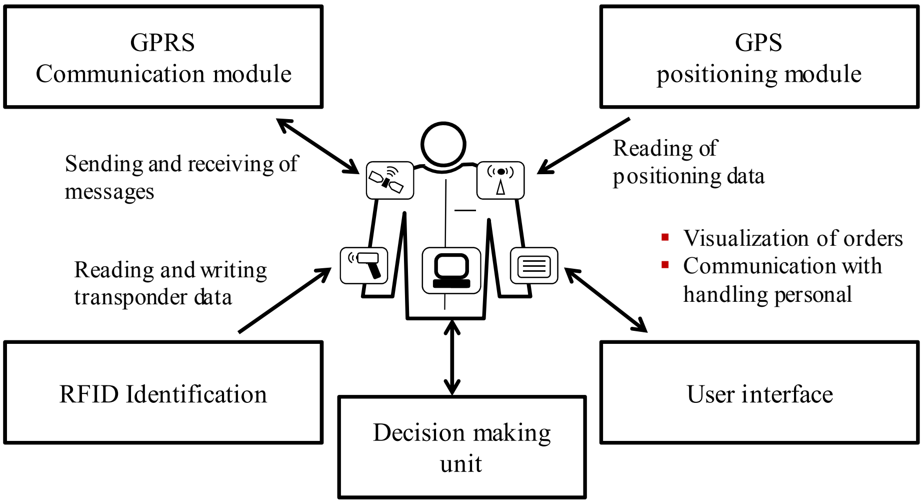


Figure 5. Components to realize autonomous control methods with wearable computing

3.2 Prototypical implementation of autonomous cooperating vehicles

A wearable computing solution fulfills the main requirements as mentioned in chapter 2 of this paper. The developed solution contains the needed required components for identification, positioning, communication as well as a user interface and a processing unit. The software concept was implemented as shown in Figure 3. Each hardware module is controlled by a separate gateway. The main control is given by the used autonomous control method which can be implemented in the decision layer.

Figure 6. Prototypical wearable computing solution for autonomous cooperating vehicles

Figure 6 illustrates the developed solution to collect the data necessary to represent the vehicles in the MAS. The main idea of this development is to equip every vehicle with an RFID transponder. The transponder contains all required data such as the Vehicle Identification Number (VIN), type, color, manufacturer for an optical identification by the terminal staff. This data should further indicate the associated vehicle orders, like washing or other technical treatment orders. The transponder data are used for the virtual representation of the vehicles. The processing unit within the wearable solution collects all required data for a decision-making by the use of the MAS as shown in Figure 4.

4. CONCLUSION

This paper focuses on the technical realization of autonomous control. A software and system related concept was introduced to represent real world logistic object in a virtual environment to enable an autonomous coordination of logistic object. As a main result of the various researches into innovative IT-based solutions the authors came to the conclusion that the technical implementation of an autonomously controlled system is already possible and makes economic sense with today's information and communication technologies.

ACKNOWLEDEMENT
This research was supported by the German Research Foundation (DFG) as part of the Collaborative Research Centre 637 "Autonomous Cooperating Logistic Processes – A Paradigm Shift and its Limitations" at the University of Bremen.

5. REFERENCES

Böse, F., Piotrowski, J., Windt, K. (2005). Selbststeuerung in der Automobil-Logistik. *Industriemanagement*, 20(4), 37-40.

Böse, F., Lampe, W., Scholz-Reiter, B. (2006). Netzwerk für Millionen Räder. FasTEr - Eine Transponderlösung macht mobil. *RFID im Blick*, Sonderausgabe RFID in Bremen, 20-23.

DIN 55022 (2007). Einrichtungen der Informationstechnik Funkstöreigenschaften - Grenzwerte und Messverfahren. Berlin: VDE Verlag.

DIN EN 55024 (2003). Einrichtungen der Informationstechnik - Störfestigkeitseigenschaften - Grenzwerte und Prüfverfahren. Berlin: VDE Verlag.

DIN61000-3-2 (2006). Elektromagnetische Verträglichkeit (EMV) Teil 3 -2: Grenzwerte - Grenzwerte für Oberschwingungsströme (Geräte-Eingangsstrom =< 16 A je Leiter)(IEC 61000-3-2:2005); Deutsche Fassung EN 61000-3-2:2006. Berlin: VDE Verlag.

DIN61000-3-3 (2006). Elektromagnetische Verträglichkeit (EMV) Teil 3 -3: Grenzwerte - Begrenzung von Spannungsänderungen, Spannungsschwankungen und Flicker in öffentlichen Niederspannungs-Versorgungsnetzen für Geräte mit einem Bemessungsstrom 16 A je Leiter, die keiner Sonderanschlussbedingung unterliegen. Berlin: VDE Verlag.

Kirisci, P., Morales, E. (2006). Der Einsatz von Wearable Computing im industriellen Kontext. *Mobilfunk: Technologien und Anwendungen*, ITG-Fachbericht Nr. 194, Osnabrück: VDE Verlag, 115-120.

Mrugala D., Ruthenbeck C., Scholz-Reiter B., Lang W. (2009). Smart Jacket as a Computing System for Automobile Warehouse Logistics. *SMART SYSTEMS INTEGRATION 2009*. European Conference & Exhibition on Integration Issues of Miniaturized Systems - MEMS, MOEMS, ICs and Electronic Components, AKA Verlag, Heidelberg, Germany, 2009, 492-495.

Scholz-Reiter, B., Freitag, M., Schmieder A. (2002). Modelling and Control of Production Systems based on Nonlinear Dynamics Theory. *Annals of the CIRP*, vol. 51, Issue 1, 375-378.

Scholz-Reiter, B., Böse, F., Teucke, M., Piotrowski, J. (2008) Integrated Planning and Control on Ro-Ro Terminals. *7th International Workshop on Modeling & Applied Simulation (MAS 2008)*, 275-280.

Windt, K., Böse, F., Philipp, T. (2005). Criteria and Application of Autonomous Cooperating Logistic Processes. *Proceedings of the 3rd International Conference on Manufacturing Research, Advances in Manufacturing Technology and Management.*

Windt, K., Hülsmann, M. (2007). Changing Paradigms in Logistics - Understanding the Shift from Conventional Control to Autonomous Cooperation and Control. *Understanding Autonomous Cooperation & Control* - The Impact of Autonomy on Management, Information, Communication, and Material Flow. Springer, Berlin, 4-16.

Windt, K., Becker, T. (2009). Applying Autonomous Control Methods in Different Logistic Processes - A Comparison by Using an Autonomous Control Application Matrix. In: *Proceedings of the 17th Mediterranean Conference on Control and Automation*. Thessaloniki, Greece, 2009.

Session A3: Transportation 1

- **Day1: Sep. 15 (Wed.)**

- **Time: 09:00 - 10:20**

- **Chair: Won Young Yun**

- **Room: Iris, 4F**

COMPARISON OF ALTERNATIVE SHIP-TO-YARD VEHICLES WITH THE CONSIDERATION OF THE BATCH PROCESS OF QUAY CRANES

Sang-Hei Choi[1], Seung Hwan Won[1*], and Chulung Lee[2]

[1]Future Strategy Research Division
Korea Maritime Institute
1652, Sangam-dong, Mapo-gu
Seoul, South Korea
e-mail: {shchoi, shwon}@kmi.re.kr

[2]Division of Information Management Engineering
Korea University
Anam-dong, Seongbuk-gu
Seoul, South Korea
e-mail: leecu@korea.ac.kr

Abstract: Container terminals around the world fiercely compete to increase their throughput and to accommodate new mega vessels. In order to increase the port throughput drastically, new quay cranes capable of batch processing are being introduced. The tandem-lift spreader, equipped with a quay crane, which can handle one to four containers simultaneously has recently been developed. Such increase in the handling capacity of quay cranes requires significant increase in the transportation capacity of ship-to-yard vehicles as well. The objective of this study is to compare the performances of three alternative configurations of ship-to-yard vehicles in a conventional container terminal environment. We assume that the yard storage for containers is horizontally configured and the quay cranes equip with tandem-lift spreaders. A discrete event simulation model for a container terminal is developed and validated. We compare the performances of the three alternatives under different cargo workloads and profiles, represented by different annual container handling volumes and different ratios of tandem mode operations, respectively. The results show that the performances of the alternative vehicle types are largely dependent on workload requirement and profile.

1. INTRODUCTION

As the size of trade between countries increased, there are rapid changes in the logistics environment concerning ports. The world container traffic in 2008 is 540 million TEUs which grew by 2.3 times compared to 230 million TEUs in 2000. It is forecasted to achieve growth rate of around the annual average of 9% by 2013. Due to this, the marine transportation industry has made *Mega-Carrier* appear through mergers and acquisitions between shipping lines to expand market dominance, and they are continuing to make enormous investments for securing mega ships over 10,000 TEUs in order to strengthen the competitiveness in shipping cost.

According to such changes in the shipping environment, large ports in the world are engaging in fierce competition for hub ports by continents in order to attract mega fleet, and this is leading to the trend of strengthening port competitiveness through the securing and operation of efficient port facilities. In other words, the world's leading ports such as Singapore, Shanghai, Hong Kong, Shenzhen, Busan, Rotterdam, and Hamburg are not only developing large-sized terminals but also investing highly productive handling equipment for the efficiency of port operation.

The handling equipment in a port generally consists of quay cranes (QCs), ship-to-yard vehicles (terminal trucks or automated guided vehicles), and yard cranes (YCs). Out of these, QCs and ship-to-yard vehicles are most closely related to ships. These are the most important factors which determine the ship turnaround time in a port.

Berthing a mega ship over 10,000 TEUs in a port requires water depth, the workable specification of QCs, and the high productivity of a terminal. Especially, despite increasing the size of ships, shipping lines tend to require the service time in the past. Therefore, ports unable to meet the trend of such customer requirements may bring about the desertion of customers.

Highly efficient operation and innovation of handling equipment in a port are being actively carried out as a part of survival strategies. Out of handling equipment in a port, the QCs have the high priority in investments. Most of ports all over the world have been operating the QCs with a *single-lift* spreader capable of processing 1 box of 20 or 40 ft container and with a *twin-lift* spreader capable of processing 2 boxes of 20 ft containers or 1 box of 40 ft container. Recently, however, the QCs with a *tandem-lift* spreader capable of processing 4 boxes of 20 ft containers or 2 boxes of 40 ft containers appeared in some ports in Netherlands, UAE, China, and South Korea. Table 1 summarizes the spreader types of QCs. In order for the QCs with a tandem-lift spreader to have high efficiency, various working method of the ship-to-yard vehicles in connection to the QC must be considered. Therefore, in preparation for the spread of the QCs with tandem-lift spreader which will emerge as the future common equipment, it is desirable to come up with the optimal composition of handling system through analysis of the impact to the terminal productivity by interaction between QCs and ship-to-yard vehicles.

Table 1. Spreader types of quay cranes (ECT Home Page; Bhimani and Sisson, 2002; Lind *et al*., 2007)

Type	Single-lift Twin- 	lift Tand 	em-lift 
Capacity	One 20ft container or one 40ft container	Two 20ft containers or one 40ft container	Four 20ft containers or two 40ft containers

The performance evaluation from the interaction between handling equipment within a terminal has been studied by various aspects.

Das and Spasovic (2003) analyzed the strategic alternatives for minimizing empty travel of straddle carriers and for minimizing waiting time of road trucks, when containers are received and delivered between road trucks and terminals. Vis and Harika (2004) analyzed the performance for each type of ship-to-yard vehicles (automated guided vehicles and automated lifting vehicles) in an automated container terminal through simulation. The performance was measured as unloading time of ships, waiting time of QCs, waiting time of vehicles, and occupancy degree of vehicles and QCs. Yang, Choi, and Ha (2004) compared the performance of ship-to-yard vehicles (automated guided vehicles and automated lifting vehicles) in an automated container terminal with simulation. They also analyzed the number of units by type of ship-to-yard vehicles satisfying the given service level and the impact it has on cycle time. Liu *et al.* (2004) analyzed the impact on the performance of terminals, using automated guided vehicles (AGVs), by yard layout (horizontal and vertical) through simulation. They used the multi attribute decision making (MADM) method to evaluate the performance of the terminal and determined the optimum number of vehicles by layout. Grunow, Günther, and Lehmann (2004) presented a priority rule based algorithm for dispatching multi-load AGVs, which can carry one 40 ft or 45 ft container or two 20 ft containers at a time, in an automated container terminal. They also compared the performance of algorithm and mathematical model by making a mixed integer linear programming (MILP) model which corresponds to the algorithm. de Koster, Le-Anh, and van der Meer (2004) compared the dispatching rules in use at three different environments (European distribution center, glass production plant, and container transshipment terminal) using simulation. They analyzed the case-specific dispatching rules and general dispatching rules (nearest-workstation-first, nearest-vehicle-first, modified-first-come-first-served, and nearest-vehicle-first with time priority) together in each environment. Duinkerken *et al.* (2006) compared the performance of three types of transportation system (multi-trailer system, AGV, and automated lifting vehicle) used in inter-terminal transport using simulation. Rule-based control system and advanced planning algorithm were included in the simulation model and the situations to be implemented in the actual terminal were compared and analyzed through various scenarios.

In this study, we analyze the impact which the type of ship-to-yard vehicles has on the terminal performance using a discrete event simulation model, when the QCs with tandem-lift spreader are introduced. We compare the performances of the three alternative ship-to-yard vehicles under different cargo workloads and profiles, represented by different annual container handling volumes and different ratios of tandem mode operations, respectively.

Section 2 introduces the tandem-lift QCs and alternative ship-to-yard vehicles. Section 3 discusses a discrete event simulation model developed and Section 4 provides the result of experiments by simulation. Finally, Section 5 presents the conclusions of this study.

2. PROBLEM DESCRIPTIONS

2.1 Tandem-Lift Quay Crane for Batch Processing

Generally, the handling system in a container terminal consists of QCs, ship-to-yard vehicles, and YCs. The productivity of the terminal is determined by the efficient connection between three types of equipment. Figure 1 shows the logistics flow in a container terminal. Import containers are delivered through quay, yard, and gate from the ship and export containers are loaded onto the ship in the opposite order.

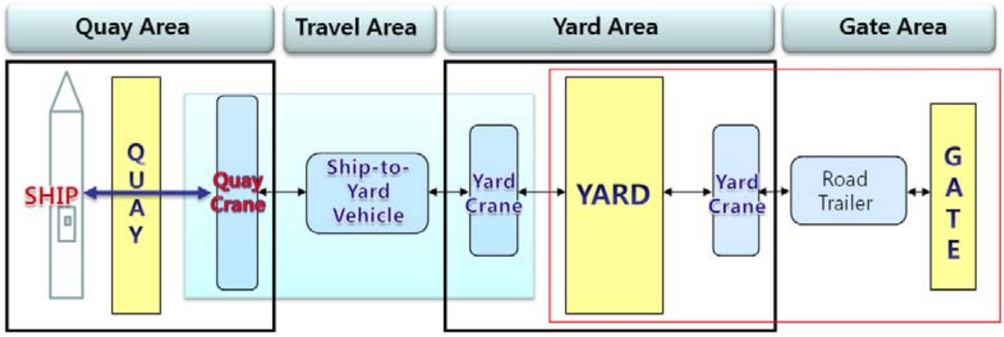


Figure 1. Container flow and equipment in a container terminal

In such logistics flow within the terminal, shipping lines would require terminal operators to reduce the ship turnaround time. Therefore, terminal operators has no choice but to satisfy the requirement level of shipping lines by investing equipment with high productivity. In this process, highly productive handling equipment such as twin-lift QCs and tandem-lift QCs appeared and there is no option but to bring about changes in the ship-to-yard vehicles in connection to these.

The theoretical productivity of tandem-lift QCs is double compared to the existing twin-lift QCs. However, when they are actually operated in a container terminal, their performance is greatly influenced by the stowage plan of ships and the types of ship-to-yard vehicles. Thus, the actual productivity of tandem-lift QCs can be determined by considering all of these factors.

In order to increase the handling performance of tandem-lift QCs, the stowage plan of ships must consider all calling ports together. The stowage plan considering only each port cannot maximize the handling performance of tandem-lift QCs in all calling ports. Because tandem-lift QCs handle multiple containers simultaneously, connecting to the ship-to-yard vehicles that can transport multiple containers is effective. The normally operated ship-to-yard vehicles can transport one 40 ft container of two 20 ft containers. When operating tandem-lift QCs, the appropriate connection to the ship-to-yard vehicles will be an important factor in determining the productivity of the terminal. Accordingly, the most appropriate ship-to-yard vehicle must be operated.

2.2 Alternative Ship-to-Yard Vehicles

Because the tandem-lift QC handles maximum 4 boxes of 20 ft containers or 2 boxes of 40 ft containers simultaneously, the ship-to-yard vehicles connected to this must accommodate multiple containers. The currently operated ship-to-yard vehicles include single-stack trailers (SSTs), double-stack trailers (DSTs), parallel dual-trailers (PDTs), and serial dual-trailers (SDTs). Table 2 shows alternative ship-to-yard vehicles. An SST has the capacity of 2 TEUs, while a DST, a PDT, and an SDT have the capacity of 4 TEUs.

Figure 2. Alternatives of ship-to-yard vehicles

2.3 Connection between Tandem-Lift Quay Cranes and Ship-to-Yard Vehicles

When the tandem-lift QC is operated, there are three factors largely influencing the terminal productivity. First, the efficiency of loading and unloading onto/from a vehicle by a QC may vary, due to the different chassis structure of ship-to-yard vehicles. Two SSTs or a PDT have the chassis structure capable of simultaneous loading and unloading. However, in the case of a DST, one spreader (one of parallel two twin-lift spreaders) and the other spreader of the tandem-lift QC must be worked on separately. In the case of an SDT, a vehicle must shift from the work for a front trailer to the work for a rear trailer, during loading and unloading by QCs.

Second, there is the dispatching problem of vehicles assigned to tandem-lift QCs. When a QC is going to process two vehicles and only one vehicle arrive at the transfer area of the QC, the QC might wait for the other vehicle. In the case of the SST, if two vehicles arrive at nearly the same time, there are no problems. But if only one vehicle arrives first, a QC might wait until the other vehicle arrives additionally. In this case, to increase the turnover of ship-to-yard vehicles, it might be more favorable to handle one vehicle first and the other vehicle later. However, such situation does not occur in the case except the SST.

Third, it is about the problems related to the ratio of tandem mode operation of QCs. Currently, the ports operating tandem-lift QCs is in test operation or are extensively used in work of empty containers. In the case of empty containers, because it is not impacted that much by the order of loading and discharging work of the ship, the ratio of tandem mode operation tends to be very high. As previously stated, the ratio of tandem mode operation is largely influenced by the stowage plan of previous and next calling ports. Therefore, most of ports currently operating tandem-lift QCs set the target ratio of tandem mode operation to within 20-30%. If the majority of ports install tandem-lift QCs in the future, the ratio of tandem mode operation is expected to be higher.

3. SIMULATION MODEL

3.1 Physical Conditions

We assume that the yard storage for containers is horizontally configured, the maximum four vessels can come alongside the terminal simultaneously, and the storage yard consists of 27 blocks. The handling equipment in the terminal consists of 16 single-trolley QCs with tandem-lift spreaders and 54 cantilever YCs with twin-lift spreader. The ship-to-yard vehicles are selected out of SSTs, DSTs, and SDTs. Because the YCs must extend their out reach when using PDTs, this study does not consider the PDT type of ship-to-yard vehicles. It is assumed that the lengths of an SST, a DST, and an SDT are 16.556 meters, 22.859 meters, and 38.151 meters, respectively.

3.2 Modeling Ship-to-Yard Vehicles

3.2.1 Process Definition

We defined the process between ship-to-yard vehicles, QCs, and YCs in order to apply the operation of each type of ship-to-yard vehicle to a simulation model. Ship-to-yard vehicles transport containers between the apron and the yard. In the case of unloading from a vessel, they move without containers from the yard to the apron and try to enter the transfer area with QCs in the apron. If possible, they receive containers from QCs and transport containers from the apron to the yard. When they arrive at the yard, they try to enter the transfer area with YCs in the yard. If possible, they deliver containers to YCs. Because ship-to-yard vehicles are passive, their receiving and delivering operation depends on the availability of QCs and YCs. Therefore, the synchronization between the arrival of ship-to-yard vehicles and the availability of QCs and YCs leads to high terminal productivity. Figure 3 shows the processes between DSTs, QCs, and YCs, in the case of unloading from a vessel.

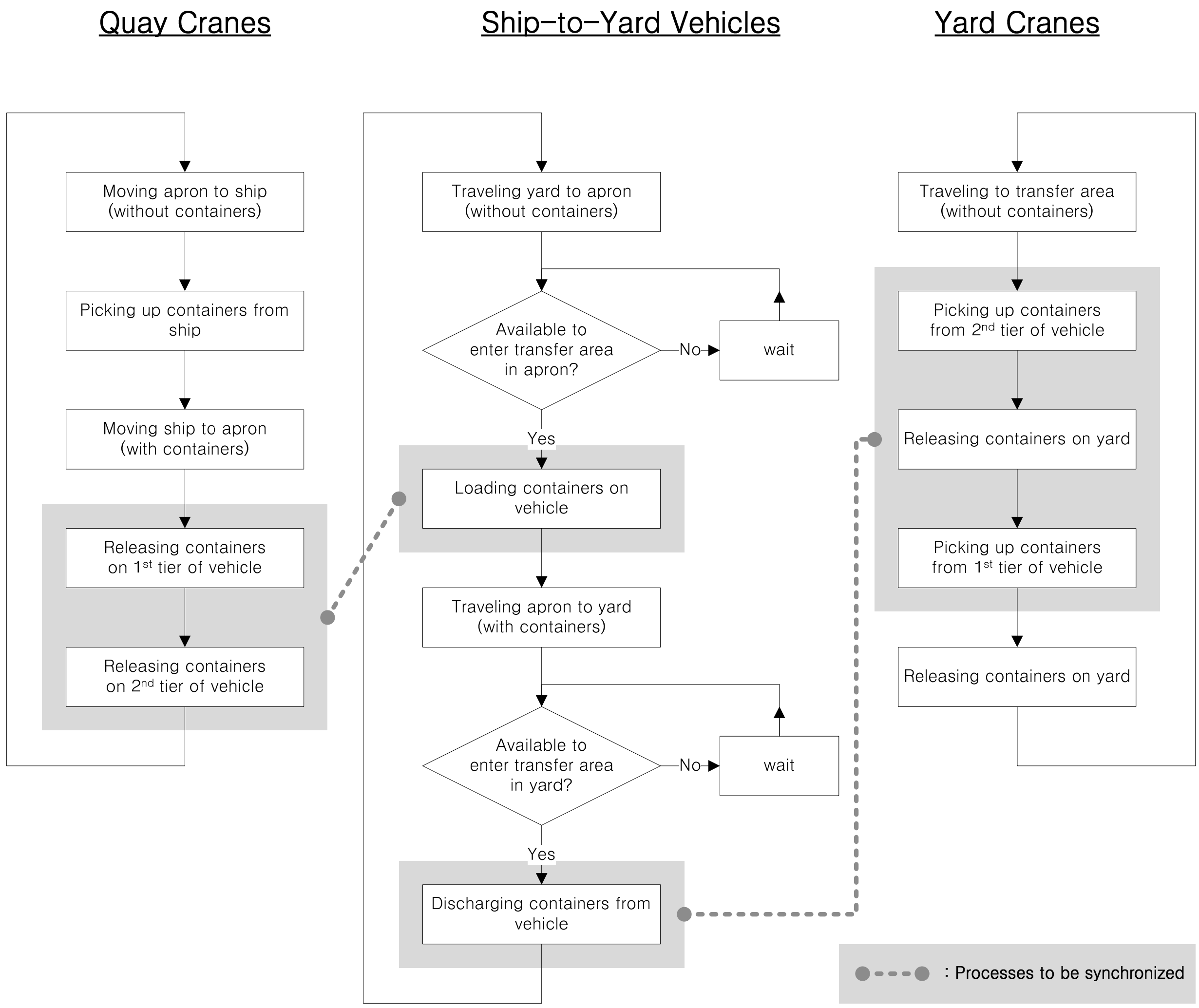
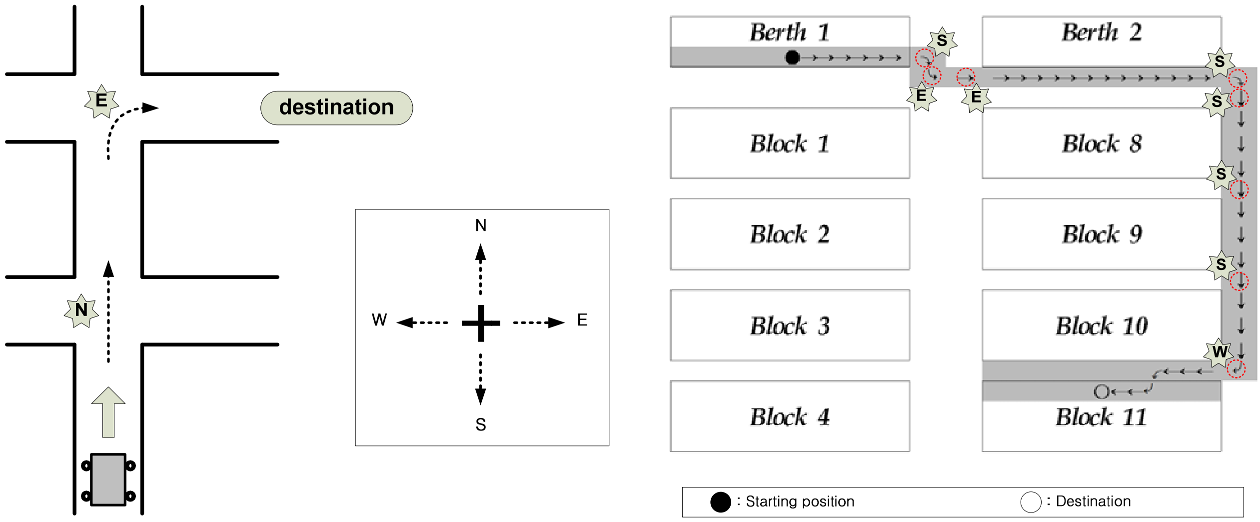


Figure 3. Flowchart for the processes of DSTs

3.2.2 <u>Routing</u>

The routing of ship-to-yard vehicles is determined by the starting position and the destination. The routing of a ship-to-yard vehicle is represented by the list of travelling directions at the intersections which a ship-to-yard vehicle passes. The direction at an intersection is one of four absolute directions which are north (N), east (E), west (W), and south (S), as shown Figure 3-(a). Figure 3-(b) illustrates the routing from berth *1* to block *11*. The routing is a list {S, E, E, S, S, S, S, W}. When a ship-to-yard vehicle passes an intersection, it withdraws, interprets, and deletes the first element of its routing list.



(a) Direction at an intersection (b) Example of routing

Figure 3. Routing of ship-to-yard vehicles

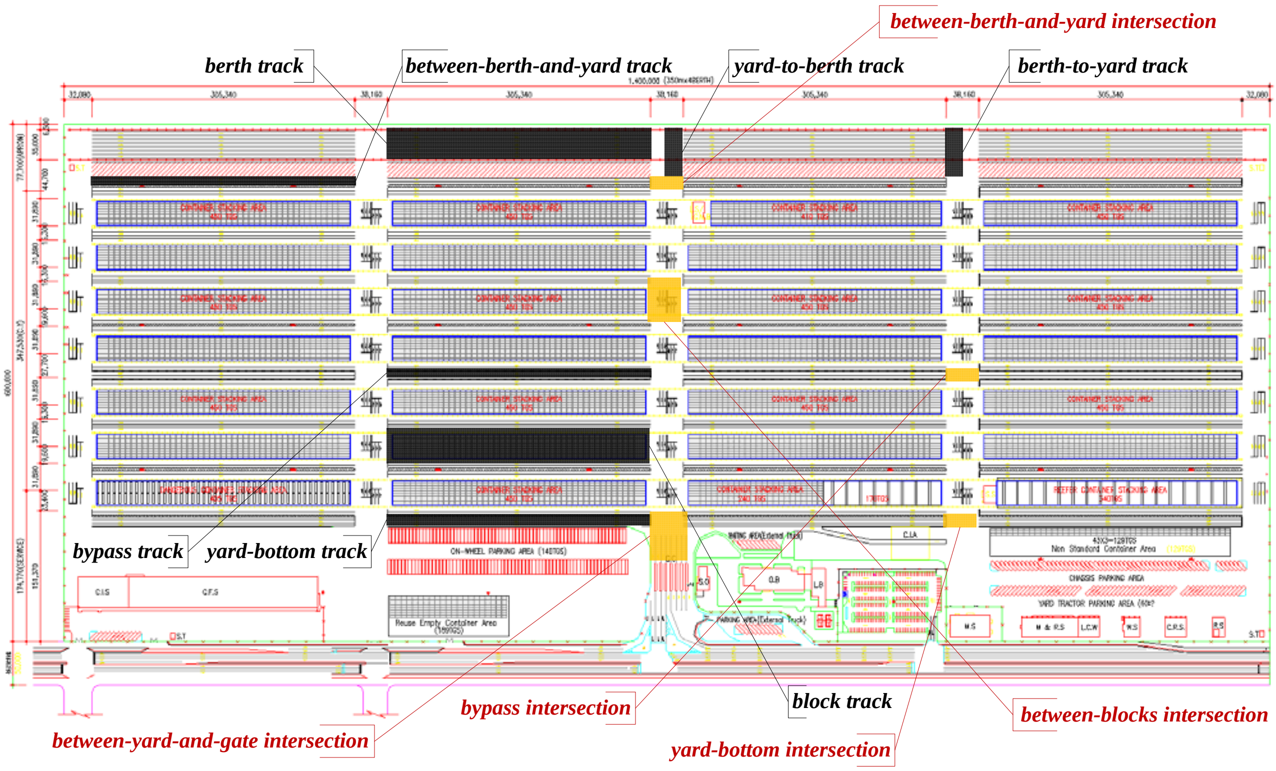


Figure 4. Definition of zones in the travel area

3.2.3 Traffic Control

The traffic of ship-to-yard vehicles is managed using the zone in the travel area. Each zone has the specific capacity. The vehicles trying to enter a zone are controlled by the remaining capacity of the zone. The travel area of ship-to-yard vehicles consists of two types of zones: one is *track-zone* and the other is *intersection-zone*. The track-zone is a set of tracks without intersection and the intersection-zone is a set of tracks with intersection, in specific areas. The track-zone is classified in to *berth track*, *block track*, *yard-to-berth track*, *berth-to-yard track*, *between-berth-and-yard track*, *bypass track*, and *yard-bottom track* according to the number and the position of tracks, similarly, the intersection-zone is classified into *between-berth-and-yard intersection*, *between-blocks intersection*, *bypass intersection*, *yard-bottom intersection*, and *between-yard-and-gate intersection*, as shown Figure 4.

3.2.4 Dispatching Rule

We assumed that a pool of ship-to-yard vehicles is dedicated to a berth. Ship-to-yard vehicles are dispatched to the QC having the most remaining tasks among the QCs assigned to the berthing vessel. This leads to the fast completion of all tasks for a vessel. When a QC simultaneously handles four 20 ft containers or two 40 ft containers, two SSTs, a DST, or an SDT are dispatched to a task. If two SSTs arrive at the transfer area with QCs in the apron within the available arrival interval (AAI), they are processed together. Otherwise, they are processed separately. However, when a QC handles one 20 ft container, two 20 ft containers, or one 40 ft containers, an SST, a DST, or an SDT are dispatched to a task. Note that the AAI can be set as an input parameter in the simulation model.

3.3 Modeling Other Equipment

3.3.1 Quay Cranes

QCs perform the gantry travel, the hoist operation, the trolley operation, the pickup from a vessel, and the release into a vessel, independently of ship-to-yard vehicles. However, their operations connected to ship-to-yard vehicles depend on the type and the arrival of ship-to-yard vehicles. Moreover, when performing the tandem mode operation, additional working times are needed to consider. The additional working times of SSTs are the arrival interval of two vehicles which are dispatched to a same QC operation but are not processed together. That is the waiting time of the QC. The additional working times of DSTs and SDTs come from two times of picking up or releasing from/onto a vehicle.

3.3.2 Yard Cranes

The gantry travel of YCs is implemented in the simulation model, while the container handling of them is considered as processing times. The processing times can be estimated by using the standard time of each operation of YCs. The values also vary according to the type of ship-to-yard vehicles and the dimension of yard blocks. We used the model proposed by Lee and Kim (2007) to estimate the standard time of YCs.

Sequencing operations of a YC is determined by nearest vehicle first served (NV) rule. A YC serves the vehicle that is located nearest it under NV rule. If vehicles are at a same distance, they are served in the order of their arrival time. If a YC has both vessel operations interfacing with ship-to-yard vehicles and receiving and delivery operations interfacing with road trucks, vessel operations have priority over receiving and delivery operations. However, the vehicles located in the bay on which the number of waiting vehicles exceeds its capacity or the vehicles exceeding the allowable waiting time are first served irrespective of distances and operation types.

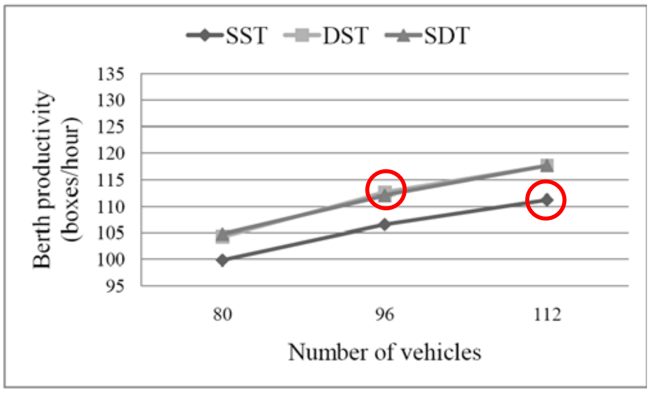
4. SIMULATION RESULT

4.1 Design of Experiment

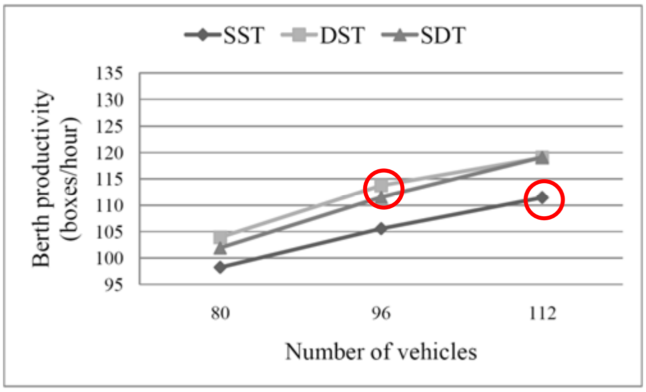
The factors of experiments consist of the annual container handling volume, the tandem ratio of QCs, the type of ship-to-yard vehicles, and the fleet size of ship-to-yard vehicles. Table 2 summarizes the conditions of experiments. Total 54 cases were generated by combining the levels of each factor. It is assumed that the type and the number of QCs and YCs are the same in all case.

Table 2. Conditions of experiments

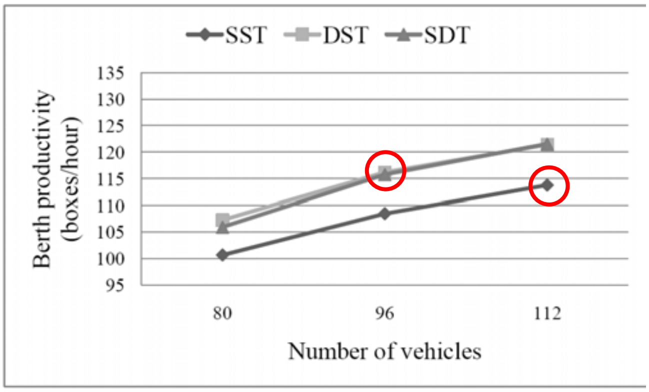
Annual container handling volume (in TEUs)	Tandem ratio of QC (ratio for unloading, ratio for loading)	Ship-to-yard vehicle	
		Type Fleet	size
2,400,000 and 2,800,000	(0.3, 0.1), (0.4, 0.2), and (0.5, 0.3)	SST, DST, and SDT	80, 96, and 112



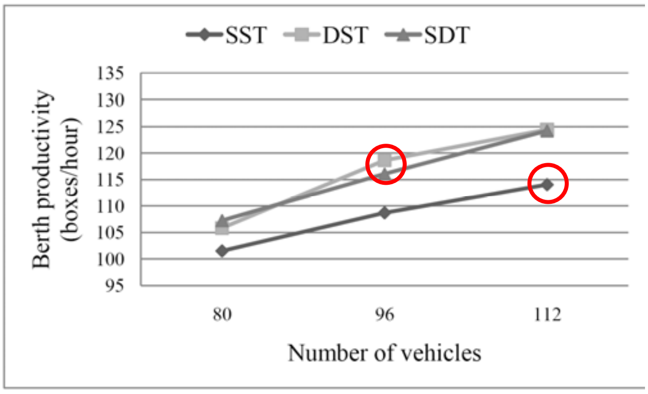
(a) Annual container volume is 2,400,000 TEUs and tandem ratios are 0.3 for unloading and 0.1 for loading



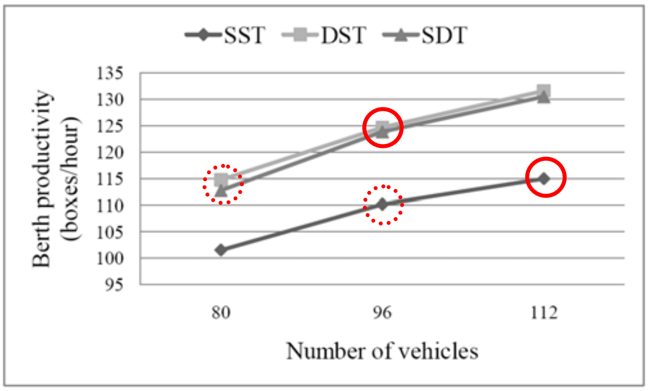
(b) Annual container volume is 2,800,000 TEUs and tandem ratios are 0.3 for unloading and 0.1 for loading



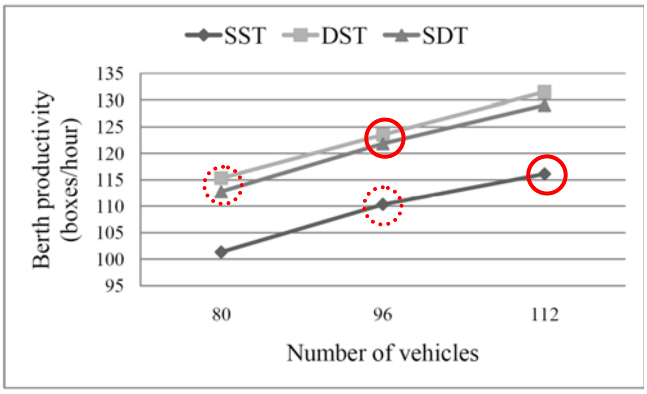
(c) Annual container volume is 2,400,000 TEUs and tandem ratios are 0.4 for unloading and 0.2 for loading



(d) Annual container volume is 2,800,000 TEUs and tandem ratios are 0.4 for unloading and 0.2 for loading



(e) Annual container volume is 2,400,000 TEUs and tandem ratios are 0.5 for unloading and 0.3 for loading



(f) Annual container volume is 2,800,000 TEUs and tandem ratios are 0.5 for unloading and 0.3 for loading

Figure 5. Average berth productivity from experiments

4.2 Result of Experiment

The average berth productivity was measured as the throughput of the terminal. Because four QCs in one berth cooperate with each other to serve a vessel, the productivity of the combined four QCs corresponds to the service capability of one berth.

Figure 5 shows the average berth productivity of all cases in experiments. From experiments, it was concluded that the productivity in the cases of DSTs or SDTs is superior to that in the cases of SSTs. The difference of productivity in the cases between of DSTs and of SDTs was insignificant. In addition, as the tandem ratios of QCs increased, the difference of productivity in the cases between of SSTs and of DSTs or SDTs increased. These were very trivial results. Because the productivity grew linearly according to the increase of the number of vehicles, it is expected that ship-to-yard vehicles were the bottleneck resources of the terminal, in the experiments. Other resources such as QCs, YCs, and the travel areas of vehicles were not critical to the performance of the terminal.

From the view of replacement of ship-to-yard vehicles, less DSTs or SDTs showed equal or higher performance than more SSTs, in some cases. If the cost of a DST or an SDT is within 117% of the cost of an SST, 112 SSTs can be replaced by 96 DSTs or SDTs (refer to the line circles in Figure 5). Moreover, if the cost of a DST or an SDT is within 120% of the cost of an SST only when the tandem ratios are 0.5 for unloading and 0.3 for loading, 96 SSTs can be replaced by 80 DSTs or SDTs (refer to the dotted circles in Figure 5).

5. CONCLUSIONS

This study discussed the terminal performance on which alternative ship-to-yard vehicles impact, when the tandem-lift QCs are introduced. The four types of ship-to-yard vehicles, SSTs, DSTs, SDTs, and PDTs, were proposed and their operations were designed. To compare the performances by each vehicle type, the discrete event simulation model was developed. The processes, routing, and the traffic control, and the dispatching rule of ship-to-yard vehicles were implemented in detail, in the simulation model. From experiments, we verified that the performance in the cases of DSTs or SDTs is superior to that in the cases of SSTs. In addition, we realized that the tandem ratios of QCs are critical to the choice of alternative ship-to-yard vehicles.

In order to optimize the fleet sizes of DSTs, SDTs and SSTs under various workload requirements and profiles, an analytical model that is capable of estimating the berth productivity is currently being developed. The proposed model describes the operations of the tandem-lift QCs and ship-to-yard vehicles using a queuing model, and computes the average throughput and the average flow-time for various system configurations. The model will then be validated using the simulation results presented in this study. Using the model, the optimal number of DSTs, SDTs and SSTs necessary to cope with given container handling requirements will be determined. The analytical model may be employed for the design of the ship-to-yard vehicle systems prior to the simulation analysis of the alternative designs.

ACKNOWLEDGEMENTS
This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0016207).

6. REFERENCES

Bhimani, A. K. and Sisson, M. (2002). Increasing Quayside Productivity. *Pan-Pacific Conference*, Pan-Pacific Business Associations.

Briskorn, D., Drexl, A., and Hartmann, S. (2006). Inventory-Based Dispatching of Automated Guided Vehicles on Container Terminals. *OR Spectrum*, 28: 611-630.

Das, S. K. and Spasovic, L. (2003). Scheduling Material Handling Vehicles in a Container Terminal. *Production Planning & Control*, 14: 623-633.

Drewry Shipping Consultants (2005). *Annual Container Market Review and Forecast - 2005/06*. Drewry, London, UK.

Drewry Shipping Consultants (2006). *Annual Container Market Review and Forecast - 2006/07*. Drewry, London, UK.

Drewry Shipping Consultants (2007). *Annual Container Market Review and Forecast - 2007/08*. Drewry, London, UK.

Drewry Shipping Consultants (2008). *Annual Container Market Review and Forecast - 2008/09*. Drewry, London, UK.

Duinkerken, M. B., Dekker, R., Kurstjens, S. T. G. L., Ottjes, J. A., and Dellaert, N. P. (2006). Comparing Transportation Systems for Inter-Terminal Transport at the Maasvlakte Container Terminals. *OR Spectrum*, 28: 469-493.

ECT (Europe Container Terminals): ECT Home Page. http://www.ect.nl

Grunow, M., Günther, H.-O., and Lehmann, M. (2004). Dispatching Multi-Load AGVs in Highly Automated Seaport Container Terminals. *OR Spectrum*, 26: 211-235.

Grunow, M., Günther, H.-O., and Lehmann, M. (2006) Strategies for Dispatching AGVs at Automated Seaport Container Terminals. *OR Spectrum*, 28: 587-610.

Hoag, M. (2008). Thailand Port and Logistics Update. *Automotive Focus Group Seminar*, Australian-Thai Chamber of Commerce.

Kim, K. H., Lee, K. M., and Hwang, H. (2003). Sequencing Delivery and Receiving Operations for Yard Cranes in Port Container Terminals. *International Journal of Production Economics*, 84: 283-292.

de Koster, R. B. M., Le-anh, T., and van der Meer, J. R. (2004). Testing and Classifying Vehicle Dispatching Rules in Three Real-World Settings. *Journal of Operations Management*, 22: 369-386.

Lee, B. K. and Kim, K. H. (2007). Cycle Time Models for Yard Cranes Considering Block Layouts in Container Terminals. *Journal of the Korean Institute of Industrial Engineers*, 33: 110-125.

Lind, D., Hsieh, J. K., and Jordan, M. A. (2007). Tandem-40 Dockside Container Cranes and Their Impact on Terminals. *PORTS Conference*, American Society of Civil Engineers.

Liu, C.-I., Jula, H., Vukadinovic, K., and Ioannou, P. (2004). Automated Guided Vehicle System for Two Container Yard Layouts. *Transportation Research Part C*, 12: 349-368.

van der Meer, R. (2000). *Operational Control of Internal Transport.* Ph. D. Thesis, Erasmus Research Institute of Management.

Song, J.-H. (2009). New Paradigm on the Lashing Process in Container Ports. *Port Technology International*, 42: 48-50.

Vis, I. F. A. and Harika, I. (2004). Comparison of Vehicle Types at an Automated Container Terminal. *OR Spectrum*, 26: 117-143.

Vis, I. F. A., de Koster, R. B. M., and Savelsbergh, M. W. P. (2005). Minimum Vehicle Fleet Size under Time-Window Constraints at a Container Terminal. *Transportation Science*, 39: 249-260.

Vis, I. F. A. (2006). A Comparative Analysis of Storage and Retrieval Equipment at a Container Terminal. *International Journal of Production Economics*, 103: 680-693.

Won, S. H. and Choi, S.-H. (2009). Design and Implementation of the Simulator for Evaluating the Performance of Container Cranes. *Journal of Intelligence and Information Systems*, 15: 119-136.

Yang, C. H., Choi, Y. S., and Ha, T. Y. (2004). Simulation-Based Performance Evaluation of Transport Vehicles at Automated Container Terminals. *OR Spectrum*, 26: 149-170.

ZPMC (Zhenhua Port Machinery Company): ZPMC Home Page. http://www.zpmc.com

PLANNING STORAGE YARD LAYOUTS OF CONTAINER TERMINALS

Joerg Wiese[1]*, Leena Suhl[1], and Natalia Kliewer[2]

[1]University of Paderborn
Germany
e-mail: wiese@dsor.de
suhl@dsor.de

[2]Freie Universität Berlin
Germany
e-mail: natalia.kliewer@fu-berlin.de

Abstract: All around the world container terminals are important nodes in the globalized transportation network. The design of the storage yard layout of a container terminal is an important impact factor for its efficiency. Several different container terminal yard layout configurations exist depending on the type of stacking equipment used. Typical yard layouts are based on terminal systems where rubber tired gantry cranes, rail mounted gantry cranes or straddle carriers are used for the stacking of containers. We define different yard layout classes for different types of stacking equipment. Based on these classes we present approaches to support the optimal design of terminal yard layouts. Currently, publications about the design of yard layouts focus mainly on terminals where rubber tired gantry cranes are used as stacking equipment. In contrast, we present an approach to design the yard layout when straddle carriers are used. Therefore, we derive different formulas to measure the expected cycle times of straddle carriers for travelling to the designated storage position and for executing the stacking operations. Results are presented for different typical terminal configurations.

SIMULATION ANALYSIS OF CRANE DEPLOYMENT STRATEGIES FOR AUTOMATED RAIL-MOUNTED-GANTRY CRANE SYSTEMS

Nils Kemme[1]*

[1]Institute for Operations Research
University of Hamburg
Germany
e-mail: nils.kemme@uni-hamburg.de

Abstract: The container yard plays as major role for the competitiveness of container terminals. One of the latest trends in container yard operations is the usage of different kinds of rail-mounted-gantry (RMG) crane systems. The performance of these systems is to a great amount dependant on the underlying crane deployment strategy which decides on the crane assignment and sequencing of transport jobs. This paper seeks to investigate which objective function and which crane deployment strategy best supports the overall yard and terminal objectives. Several crane deployment strategies are evaluated for different crane systems and yard block layouts by means of a discrete event simulation model which reproduces the stochastic and highly dynamic environment of a seaport container terminal.

Session A4: VRP

·**Day1: Sep. 15 (Wed.)**

·**Time: 09:00 - 10:20**

·**Chair: Shiwoo Lee**

·**Room: Grand Ballroom, 5F**

LOGMS

The 1st International Conference on Logistics and Maritime Systems

September 15th–17th, 2010 Busan, Korea

Lin

ITERATED LOCAL SEARCH FOR MULTI-CRITERIA AND MULTI-STAGE LOCATION ALLOCATION PROBLEM

Natthapong Nanthasamroeng[1*] and Rapeepan Pitakaso[2]

[1,2]Department of Industrial Engineering
Faculty of Engineering, Ubon Rachathani University
Warinchamrab,Ubon Ratchathani THAILAND 34190
e-mail:nnanthas@yahoo.com

Abstract: This paper presents an iterated local search algorithm for solving a multi-criteria and multi-stage location allocation problem with an application in real case of ethanol plant in Thailand. The purpose of formulated model is to find an optimal solution for locating the ethanol plant for utilization of excess bagasses from 13 sugar mills and delivered ethanol product to 4 blending centers. Three criterions are considered including: (1) minimize an operating cost and transportation cost, (2) minimize an environmental impact focus on global warming potential (GWP) and (3) minimize a societal risk from ethanol transportation.

1. INTRODUCTION

The severity of energy crisis was increasing since the unstoppable raising of crude oil price in global market. Several alternative energies for vehicles were considered such as natural gas (NGV), fuel cells, hybrid energy and ethanol blended gasoline or gasohol. Nowadays, gasohol E10, which mixed 10 percent of bio-ethanol to gasoline, is the most popular alternative fuel for automobile user. Moreover, several car manufacturers are developing new engine which can use a new gasohol which mixed ethanol up to 85 percent to gasoline. There are several type of material for producing ethanol, but most of ethanol production are from agricultural or agro-industries products such as cassava, sugarcane, or barley. The beginning of E10 production starts at farm or sugar mill. Materials were sent from their source to produce ethanol at the plant. Then, the ethanol, a kind of hazardous material, will be transported to gasohol blending centre and mix to gasoline.

In this research, we presented a multi-criteria model to locate the bagasse ethanol plant. The aim of our multi-criteria, multi-stage production system is to decide where to open ethanol plant. Three criterions are considered including; (1) minimize total cost, (2) minimize an environment impact focus on global warming potential (GWPs) and, (3) minimize a societal risk from ethanol production and transportation.

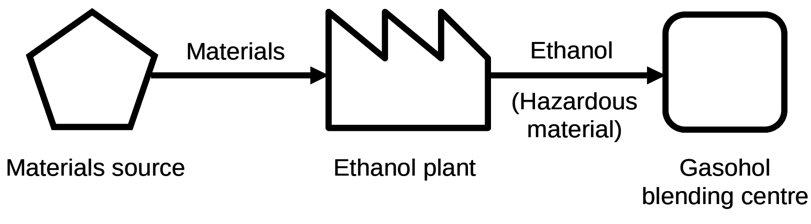


Figure 1. Multi-stage models for ethanol production

1.1 Literature Review

Facility location analysis plays an important role in many perspectives such as selection of the best location for treatment plant for hazardous waste, selection the warehouse or distribution centre in the supply chain (ReVelle and Eiselt, 2007).

Production of E10 is 2-echelon production system which similar to the two-level location-routing problem for newspapers production and delivery as Jacobsen and Madsen (1980) production system model.

Management of hazardous materials (HAZMAT) is the one in a little number of papers in a field of multi-objective location analysis. There are more than 3,300 substances was characterized as HAZMAT; this items comprises ignitability, flammable, corrosiveness, reactivity and toxicity. Various risk measures are used in HAZMAT papers such as societal risk and population exposure. Societal risk is calculated by multiply the probability of HAZMAT accident occurrence by the consequences of that accident. On the other hand, the population exposure is the number of habitant exposed to HAZMAT (Alumar and Kara, 2007)

Buddadee et al. (2008) conducted LCIA to quantify the global warming potentials, which is one type of simplified index based upon radiative properties that can be used to estimate the potential future impacts of emissions of different greenhouse gases (GHGs) upon the climate system in a relative sense, for the utilization of the excess bagasses generated from sugar industry by 2 schemes. In the first scheme, excess bagasses were used as boiler's biomass fuel for producing high pressure steam and subsequently generating electricity for internal uses. For the second scheme, the excess bagasses were sent to ethanol plants for ethanol production and conveyed to blending centre for producing gasohol E10.

Focus on the second scheme, there are 3 GWPs components comprise of GWPs due to transportation excess bagasses from sugar mill to ethanol plant, GWPs due to ethanol production and, the offset GWPs due to utilization of produced ethanol as gasohol E10 fuel in gasoline vehicle.

Nanthasamroeng et al. (2008) had studied the location allocation problem for 6 ethanol plant in northeastern of Thailand. Production and transportation framework was similar to this study. The study shown that prioritized of each objective lead to difference of locations and number of plants opened. However, all six potential plants in the study were supposed by using center of gravity techniques.

Again, Nanthasamroeng et al. (2009) had improved their studied by supposed the potential plant at all 13 sugar mills.The computational result from LINGO V.11 shown that optimal solution for ethanol plant location was to open plant number 03 and 13 which would be calculated objective value 164.9 Million Baht.

Iterated local search (ILS) is a metaheuristic algorithm that have 2 main characteristics including (1) there must be a single chain that is being followed and, (2) the search for better solutions occurs in a reduced space defined by the output of a black-box heuristic. (Lourenco et al, 1999). Moreover, Stutzle (2003) stated that ILS has a simplicity of implementation and state-of-art results.

The purpose of this study is to introduce an ILS algorithm for multi-criteria and multi-stage location allocation problem with an application in real case of ethanol plant in Thailand.

2. RESEARCH FRAMEWORK AND METHODOLOGY

Figure 2. Research Framework

A formulation of mathematical model which improved from Nanthasamroeng et al. (2008).Then, the heuristics algorithm was designed and development. After that the algorithm was verified with real data that was collected in Northeastern of Thailand.

2.1 Data and Information Collection

The data used in this model are divided into three sets including; (1) economic data, (2) environmental data and, (3) risk data. The economic and environmental data are calculated from several factor based on LCIA and Life Cycle Cost (LCC) method. Sources of information came from several channels such as Truck and trailer suppliers, Japan Transportation Cooperation Association, SimaPro V5.1 (LCA Software), Emergency Response Guidebook, and government sector of Thailand. Sources of data and information were described in Table 1

Table 1. Sources of data

Process	Data & Information Source	s of information
Transportation	• Transportation cost • Vehicle maintenance cost • Vehicle fuel cost • Truck driver cost • Vehicle capacity • GHGs emission from vehicles • Societal risk from ethanol leakage	- Truck and trailer supplier - Japan Transport Cooperation Association, 2004 - PTT, 2006 -SimaPro V5.1 (LCA software) - Emergency Response Guidebook, 2005 - National Statistic Office, 2008
Ethanol production	• Plant installation cost • GHGs emission from ethanol production process	- Kadam,2002 - Wooley et al.,1999 - Aden et al.,2002 - SimaPro V5.1(LCA software)

(Source: Adopt from Buddadee, 2008)

2.2 Data of sugar mill location and bagasses amount in Northeastern of Thailand

An existing sugar industry in the northeastern region of Thailand is chosen for illustration of this mathematical model. There are 19 provinces in this region with population ranges from 338,700 to 2,555,346 people (National Statistical Office of Thailand, 2008). But, there are 13 sugar mill located in only 7 provinces including Burirum, Mukdahan, Udonthanee, Kalasin, Khonkaen, Chaiyabhum and, Nakornratchasrima as shown in Figure 3.

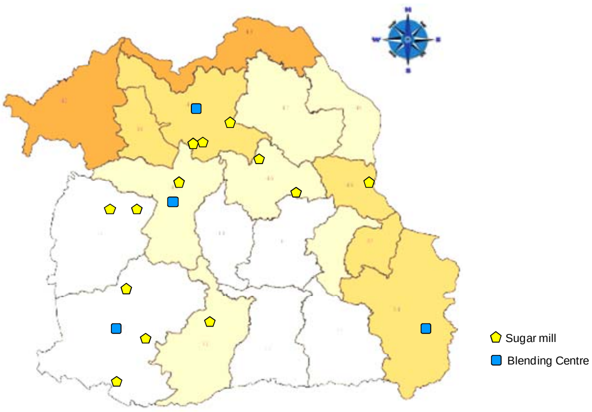


Figure 3. Location of sugar mill and blending centre in northeastern of Thailand.

The excess bagasse is considered as the material for ethanol production in this case because of environmental issue. Nowadays, excess bagasses were used as a fuel for electric generation which emitted 582,177 tons of CO_2 equivalents per year (Buddadee et al., 2008). Based on the production year 2002-2003, the excess bagasse from each sugar mill has been calculated in Table 2.

Table 2. Amount of excess bagasse in each sugar mill

Plant Code	Factory	Amount of excess bagasse (Tons/year)
01	Burirum sugar mill	36,408
02	Sahareong sugar mill	34,150
03	Reum-Udom sugar mill	68,129
04	Kasetphon sugar mill	52,631
05	Kumpawapee sugar mill	52,303
06	Khon-kaen sugar mill	87,092
07	Mitrphuwieng sugar mill	90,239
08	Roumkasettrakorn Utsahakam sugar mill	104,983
09	Utsahakamkorat sugar mill	89,330

Plant Code	Factory	Amount of excess bagasse (Tons/year)
10	Angwean (ratchasrima) sugar mill	89,952
11	N.Y. sugar mill	61,628
12	Utsahakamnamtan-Esarn sugar mill	36,663
13	Mitrkalasin sugar mill	61,259
	Total	864,406

2.3 Mathematical Model Formulation

Mathematical Model was formulated and adopt from Nanthasamroeng et al. (2008), Buddadee et al. (2007), Alumar and Kara (2005) and Wu,T.-H. (2002)

The following indices, parameters and, decision variables are used in the mathematical model:

a. Parameters and Decision Variables:

m cost of material per unit
θ cost of transportation one unit of material
a cost of transportation one unit of ethanol
c cost of base case ethanol plant
s size of base case ethanol plant
n_{ij} amount of material transported through link (i,j)
e_{jk} amount of ethanol transported through link (j,k)
o_i amount of material available in generation node i
λ_j amount of ethanol produced in ethanol plant j
d_{ij} distance between material generation node and
r_{jk} potential ethanol plant distance between potential ethanol plant and blending centre
t emission factor for the transportation of material / ethanol through link (i,j) / (j,k)
b emission factor for the diesel consumption used in transportation of material / ethanol through link (i,j) / (j,k)
f emission factor of chemical used in production of ethanol
γ emission factor due to production of diesel used in production of ethanol
g emission factor of CO_2 due to production of ethanol
h emission factor of CH_4 due to production of ethanol
δ emission factor of electric generation by using material as a fuel for the production of ethanol
v offset emission factor of E10 for the utilization of produced ethanol as E10 fuel
w offset emission factor of gasoline for the utilization of produced ethanol as E10 fuel
p offset emission factor of production of conventional gasoline for the utilization of produced ethanol as E10 fuel
$POPU$ number of population in the bandwidth for ethanol plant (j)
POP_{jk} number of population in the bandwidth for ethanol along link (j,k)
ϕ production efficiency factor for utilization of material from biomass to ethanol
N maximum number of ethanol plant
CC Carbon credit price
SC Spill compensation cost

b. Objective Function

$$\text{Minimize} \sum_{j=1}^{J}\sum_{i=1}^{I} mn_{ij}y_{ij} + \sum_{j=1}^{J}\sum_{i=1}^{I} \theta d_{ij}n_{ij}y_{ij} + c\left[\frac{\sum_{j=1}^{J}\sum_{i=1}^{I} n_{ij}y_{ij}}{s}\right]^{\exp} + \sum_{k=1}^{K}\sum_{j=1}^{J} ar_{jk}e_{jk}x_{jk} + cc\Big(\sum_{i\in I\cup J^b}\sum_{j\in I\cup J^b}\sum_{k\in K^b}(t+b)C_{ij}{}^{b}d_j^b x_{ijk}{}^{b} + \tag{1}$$

$$\sum_{i\in I\cup J^b}\sum_{j\in I\cup J^b}\sum_{k\in K^b}(f+\gamma+g+h+\delta+v+w)d_j^b x_{ijk}{}^{b} + \sum_{i\in I\cup J^e}\sum_{j\in I\cup J^e}\sum_{k\in K^e}(t+b)C_{ij}{}^{e}d_j^e x_{ijk}{}^{e} + \sum_{i\in I\cup J^e}\sum_{j\in I\cup J^e}\sum_{k\in K^e} pC_{ij}{}^{e}d_j^e x_{ijk}{}^{e}\Big) +$$

$$SC\left(\sum_{k=1}^{K}\sum_{j=1}^{J} POP_{jk}e_{jk} + \sum_{j=1}^{J} POPU_j z_j\right)$$

c. Constraints

Subject to;

$$\sum_{j=1}^{J} n_{ij}y_{ij} = o_i \qquad \forall i \tag{2}$$

$$y_{ij} = \begin{cases} 1 \text{ if generation node } i \text{ send material to ethanol plant } j \\ 0 \text{ ,otherwise} \end{cases} \qquad \forall i \forall j \tag{3}$$

$$\sum_{j=1}^{J} y_{ij} \geq 1 \qquad \forall i \tag{4}$$

$$z_j = \begin{cases} 1 \text{ ,if ethanol plant } j \text{ is open} \\ 0 \text{ ,otherwise} \end{cases} \qquad \forall j \tag{5}$$

$$z_j = y_{ij} \qquad \forall i \forall j \tag{6}$$

$$\sum_{j=1}^{J} z_j \leq N \tag{7}$$

$$\sum_{k=1}^{K} e_{jk}x_{jk} = \lambda_j \qquad \forall j \tag{8}$$

$$\lambda_j \leq \phi n_{ij} \qquad \forall i \forall j \tag{9}$$

$$x_{jk} = \begin{cases} 1 \text{ if ethanol plant } j \text{ send ethanol to blending centre k} \\ 0 \text{ otherwise} \end{cases} \qquad \forall j \forall k \tag{10}$$

$$\sum_{k=1}^{K} x_{jk} \geq 1 \qquad \forall j \tag{11}$$

$$z_j = x_{jk} \qquad \forall j \forall k \tag{12}$$

The first constraint is the mass balance constraint for material. This constraint ensures that all materials available in generation node are adequate for transportation to ethanol plant. The second and third constraint is indicate the 0-1 variable representing the available of material from source *i* to ethanol plant *j*. The forth constraint also indicate the 0-1 variable representing presence or absence of ethanol plant *j*. The fifth constraint ensured that the opened ethanol plant will have materials supplied from any sugar mills. The sixth constraint is limitation constraint for number of ethanol plant. The seventh constraint is the mass balance constraint for ethanol. The eight constraints represent the production efficiency of ethanol plant. The ninth and tenth constraint is indicating 0-1 variable represent the presence or absence of ethanol delivered from ethanol plant to blending centre.

2.4 Decoding algorithm

In this research, we use a decoding method called "Randomized Binary Selection" or RBS. RBS start with array *n* column; while *n* represent number of customer. At first step, all columns will be filled with random number 0-1. The first column represents open criteria of all plants. If a number in the column less than number in the first column, the number in this column will be changed to 0.On the other hands, the number in column that have higher value than the first column will be changed to 1. In this method, 0 represent "close" and 1 represent "open".

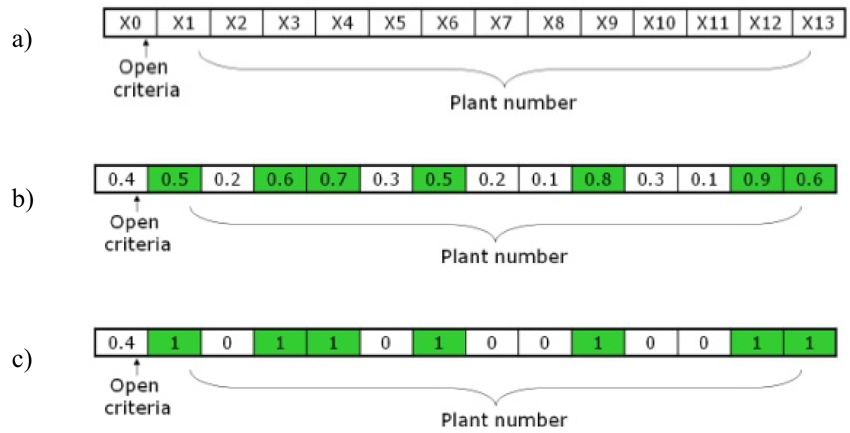


Figure 4. Illustrate of RBS algorithm

As described above, we can write pseudo-code for RBS as shown in Figure 5.

Procedure RBS
Generate 1 dimension array with n column
Random number in each column (x_0-x_n)
for $i = 1$, $i < n$ do;
if $x_n < x_0$ then $x_n = 0$, else = 1;
end for loop;
end

Figure 5. A Pseudo-code for RBS

2.5 Iterated local search algorithm

Iterated local search was officially defined by Lourenco, Martin and Stutzle (2002). ILS consist of 4 components including (1) Generate initial solution (S_0), (2) Perturbation that modified current solution (S*) to inception solution (S'), (3) Local search that improve the solution to (S'') and (4) Acceptance criterion that decide the application of perturbation(S*). A procedural view of ILS is shown in Figure 6.

s_0 ← Generate Initial Solution ( )
s^* ← Local Search (s_0)
repeat
s' ← Perturbation (s^*)
s'' ← Local Search (s')
s^* ← Acceptance Criterion (s^*, s'')
until terminate criterion met

Figure 6. A Pseudo-code for ILS procedure

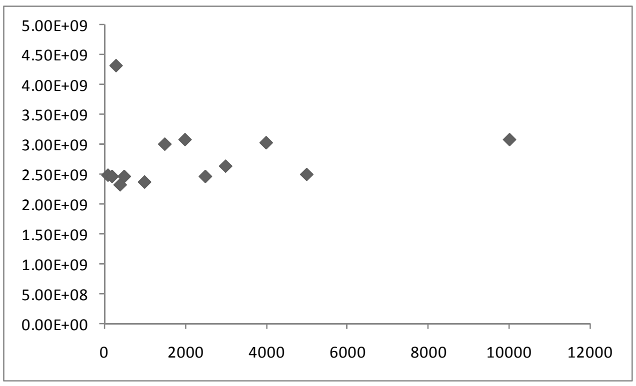
We proposed 2 algorithms, ILS-SS-SBM and ILS-SS-SBMRW, to solve our MCMSLAP and compare the result from both algorithms. The ILS-SS-SBM represents "Iterated local search with simply swap and single bit mutation" and the ILS-SS-RW represents "Iterated local search with simply swap and single bit mutation plus random walk". Main features of both algorithms are local search and perturbation process. In local search process, we use simple swap column x_n to x_{n+1}. And for perturbation process, in ILS-SS-SBM, we use single bit mutation to change "0" to "1" in selected column. The ILS-SS-SBMRW was improved to jump away from current search space if its objective value is not improved with 5 iterations.

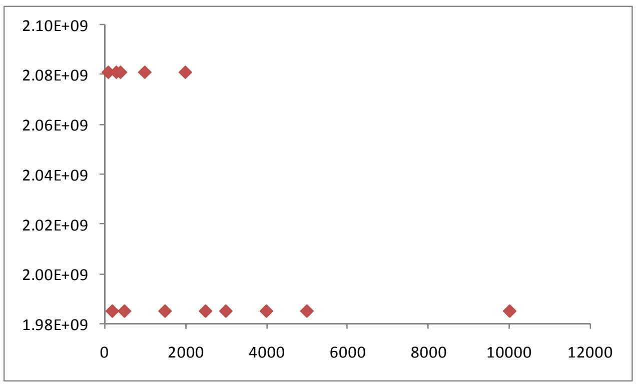
3. COMPUTATIONAL RESULT

Computational experiments were carried out on a Intel Core2Duo CPU 2.00GHz with 1.99GB of RAM. We compared result from both ILS-SS-SBM and ILS-SS-SBMRW. Both algorithms were coded in Visual C++.
We test our algorithms with real case data and define number of iteration from 100 to 10,000 iterations. The computational result was shown in table 3.

Table 3. Comparison of computational result

Iteration	**ILS-SS-SBM**			**ILS-SS-SBMRW**		
	Plant Open	**Objective Value**	**Computational Time (ms)**	**Plant Open**	**Objective Value**	**Computational Time (ms)**
100 05	,11,13	2.4835e9	32	02	2.0809e9	125
200 0	4,05,13	2.4606e9	31	01,02	1.9851e9	234
300 Al	l plants	4.3119e9	47	02	2.0809e9	281
400 0	2,13	2.3209e9	63	02	2.0809e9	359
500 0	4,05,13	2.4606e9	78	01,02	1.9851e9	406
1000 0	5,13	2.3688e9	94	02	2.0809e9	656
1500 0	1,02,03,05,09,13	3.0008e9	140	01,02	1.9851e9	828
2000 0	1,03,05,09,11,13	3.0764e9	78	02	2.0809e9	1,188
2500 0	4,05,13	2.4606e9	125	01,02	1.9851e9	1,297
3000 0	4,05,11,13	2.6335e9	78	01,02	1.9851e9	1,546
4000 0	2,04,05,06,11,13	3.0246e9	219	01,02	1.9851e9	2,234
5000 0	2,06,13	2.4947e9	219	01,02	1.9851e9	2,437
10000 0	1,03,05,09,11,13	3.0764e9	375	01,02	1.9851e9	4,844





a) ILS-SS-SBM b) ILS-SS-SBMRW

Figure 7. Comparison of ILS-SS-SBM and ILS-SS-SBMRW

4. CONCLUSIONS

This paper presents an iterated local search algorithm for solving a multi-criteria and multi-stage location allocation problem with an application in real case of ethanol plant in Thailand. The purpose of formulated model is to find an optimal solution for locating the ethanol plant for utilization of excess bagasses from 13 sugar mills and delivered ethanol product to 4 blending centers. Three criterions are considered including: (1) minimize an operating cost and transportation cost, (2) minimize an environmental impact focus on global warming potential (GWP) and (3) minimize a societal risk from ethanol transportation. We introduce 2 ILS algorithms for MCMSLAP and compare a result of both algorithms. The computation result shown that ILS-SS-SBM has a computational time less than ILS-SS-SBMRW. However, the ILS-SS-SBMRW can find the minimum cost of this problem.

For the future research, the ILS key success factors including (1) Quality of initial solution, (2) Strength of perturbation, and (3) Characteristics of local search should be considered.

5. REFERENCES

Buddadee, Bancha et al., (2007). The development of multi-objective optimization model for excess bagasse utilization: A case study for Thailand. *Environmental Impact Assessment Review*, 28 ; 380-391

Caballero, Rafael et al., (2007). Solving a multiobjective location routing problem with a metaheuristic based on tabu search. Application to a real case in Andalusia. *European Journal of Operation Research,*177; 1751-1763.

Giannikos, Ioannis, (1998). A multiobjective programming model for location treatment sites and routing hazardous wastes. *European Journal of Operation Research***,**104; 333-342.

Kara, Bahar Y. and Alumar, Sibel, (2007).A new model for the hazardous waste location-routing problem. *Computers & Operations Research* 34; 1406-1423.

Laurent, Benoit and Hao, Jin-Kao. (2009). Iterated local search for the multiple depot vehicle scheduling problem. *Computers & Industrial Engineering*, 57:277-286

Lin, C.Y.K. and Kwok, R.C.W., (2006). Multi-objective metaheuristics for a location-routing problem with multiple use of vehicles on real data and simulated data. *European Journal of Operation Research* 175; 1833-1849.

Nagy, Gabor and Salhi, Said, (2007). Location-routing: Issue, models and methods. *European Journal of Operation Research* 177; 649-672.

Nanthasamroeng,N., Pitakaso,R. and, Buddadee,B. (2008). A multi-objective model for multi-echelon location problem: Application in ethanol plant location analysis in Thailand. *Proceeding of papers, Asia conference on Intelligent Manufacturing & Logisitcs Systems*, Waseda University, Kitakyushu, Japan, 25-27 February 2008.

A HYBRID SA ALGORITHM FOR INLAND CONTAINER TRANSPORTATION

Won-Young Yun[1], and Wen-Fei Wang[2]

[1]Department of Industrial Engineering
Pusan National University
30 Jangjeon-Dong, Geumjeong-Gu
Busan 609-735, South Korea
e-mail: wonyun@pusan.ac.kr

[2]Department of Logistics Information Technology
Pusan National University
30 Jangjeon-Dong, Geumjeong-Gu
Busan 609-735, South Korea
e-mail: woyaojianchi@pusan.ac.kr

Abstract: Inland container transportation refers to container movements between customer locations, container terminals, and depots in a local area. In this paper, containers are defined as four types according to the orientations (inbound vs. outbound containers) and the container states (full vs. empty containers). And they can be delivered not only by truck but also by rail if time-windows of them can be satisfied. A mathematical model is developed graphically and a hybrid SA algorithm is proposed. The performance is addressed by numerical examples.
Keywords: Inland container transportation; time windows; traveling salesman problem (TSP); truck and rail transportation; hybrid SA (Simulated annealing)

1. Introduction

In the inland container transportation, containers are transported between container terminals, depots and customer's places by trucks and trains (R. Zhang, W.Y. Yun and I.K. Moon, 2009). The four types of containers are considered: inbound full, outbound full, inbound empty and outbound empty containers. The transportation flow depends on the type of containers. For an example, outbound freight transportation by containers can be briefly described as follows: First, a truck is assigned to carry an empty container to a customer's place and freight is loaded.. Next, the full container is delivered to a terminal directly or a rail station to transfer to the terminal. Finally, the freight is transferred to another terminal by vessel. We should determine a truck schedule to transport containers with time windows.

Inland container transportation problem with time windows may be classified into traveling salesman problem (TSP) or vehicle routing problem (VRP). There are a number of papers in which various methods to find optimal or good solutions are proposed in this area. Wen and Zhou (2007) developed a GA to solve a container vehicle routing problem in a local area. Jula *et al*. (2005) formulate truck container transportation problems by trucks with time constraints as an asymmetric multi-Traveling Salesman Problem with Time Windows (m-TSPTW). They applied DP/GA hybrid (dynamic programming and genetic algorithm) for large size problems. Zhang et al. (2009) addressed a similar problem, a graph model is built up, and a cluster method and RTS (reactive tabu search) are proposed. Liu and He (2007) decomposed a vehicle routing problem into several sub-problems according to vehicle- customer assigning structure and a new tabu search algorithm is applied to each sub-problem respectively.

In this paper, we consider trucking and rail transportation modes together to transport four types of containers in inland transportation cases. First, we build a mathematical formula by a graphical model and then propose a heuristic method. Finally, we investigate the performance of the proposed method by examples.

The remainder of this paper is organized as follows. The inland container transportation problem is described in Section 2. A graph model and a mathematical model are given in Section 3. A hybrid SA algorithm is developed to solve the problem in Section 4, and tested with a number of randomly generated examples in Section 5. Finally, Section 7 concludes this paper.

2. The inland container transportation problem

2.1 Problem definition

In this paper, a company is considered that the company possesses containers and trucks and offers container delivery services. They use trucks and trains to move containers between the terminal and customer's places. There is also a train station to transport containers from and to a container terminal. In order to not only perform good services but also remain competitive, rail transportation can be adopted to save costs. A more detailed analysis on it will be performed in the Section 3. To minimize total transportation cost is the objective in this paper.

There are four kinds of containers: inbound full, inbound empty, outbound full, outbound empty. The situation of inbound containers means they should be delivered from the terminal to costumer locations or the depot, and the situation of outbound containers means they should be delivered to the terminal.

Further assumptions and the parameters are stated in following two subsections.

2.2 Assumption

1) Only one terminal and one depot exist in the system. The terminal is a maritime terminal. An outbound container means the transporting direction of the container is from a customer or the depot to the terminal. An inbound container is exactly a reverse process. The depot has enough empty containers and trucks. The depot can be visited at any time.
2) The rail line connects the terminal and one station. The station can be used to transport containers. That is, a truck can deliver an outbound container to the station or an inbound container from the station instead of going to the terminal.
3) All containers have the same size (40ft).
4) All the trucks are identical. One truck carries no more than one container at any time. They are initially located at the depot and finally return to the depot.
5) The traveling time between any two locations is given as a constant.
6) The costs of loading/unloading and packing/unpacking are not considered, while times of them should be taken into account.
7) Each inbound or outbound full container has an origin time-window and a destination time-window. An inbound empty container only has an origin time-window because of no determinate destination. And an outbound empty container only has a destination time-window because of not determinate origin.

2.3 Parameters

C_t : Unit cost per unit time under a truck.

C_s : Unit cost per container by train.

t_{ij} : Travelling time from location *i* to location *j* by truck, *i*=T or *j*=T if the corresponding vertex is the terminal, *i*=0 or *j*=0 if the corresponding vertex is the depot, *i*=S or *j*=S if that is the station; $t_{ij} = t_{ji}$.

RT : Traveling time of the train between the station and terminal.

$[\tau_{Ai}, \tau_{Bi}]$: Time window of the origin of container *i*.

$[\tau_{Ci}, \tau_{Di}]$: Time window of the destination of container *i*.

$tlus_i$: Loading/unloading time for container *i* at the station.

$tlut_i$: Loading/unloading time for container *i* at the terminal.

$tlud_0$: Loading/unloading time for an empty container at the depot.

$tluc_i$: Loading/unloading time for container *i* at its customer location.

$tpuc_i$: Packing/unpacking time for container *i* at its customer location.

3. Problem modeling

In this section, a graphical model should be built up first; then, the problem is formulated as a mathematical model based on the graphical model.

3.1 Graphical modeling

In the Section 2, four kinds of containers are mentioned. The transportation of these containers can form a network. This network mainly focuses on activities by trucks. When the detailed observations are given to the structure of the network, it is found that there are two main parts. One is activities related to the service a truck performs while the other is activities related to the process a truck leaves from last location and gets ready for the next service. That is, the activities related to a service are treated as an entire entity and other activities are treated as connections of these entities. The entities are named vertices and the connections are named arcs. Then the graphical model can be proposed below.

Let $G=(V,A,T,C)$ indicate this graphical model, where *V* means vertices, *A* denotes arcs, *T* is the time needed for each vertex and arc, and *C* stands for transportation costs for each vertex and arc. Detailed characteristics on vertices and arcs are discussed in the following subsections. More detailed notation is shown below.

V : Set of vertices, including containers and depot, excluding the station and terminal.

$V_D=\{0\}$: Depot vertex.

V_C : Container vertices.

V_{IF} : Inbound full container vertices.

V_{IE} : Inbound empty container vertices.

V_{OF} : Outbound full container vertices.

V_{OE} : Outbound empty container vertices.

$V=V_D\bigcup V_C$, $V_C=V_{IF}\bigcup V_{IE}\bigcup V_{OF}\bigcup V_{OE}$

$A=\{(i,j)\mid i,j\in V;i\neq j\}$: Set of arcs.

3.1.1 Vertices and time associated with vertices

There are two types of vertices according the graphical model definition: one depot vertex and many container vertices. The depot vertex is a special vertex, for it involves no actual activities. It is only the start and return vertex for all trucks. Compared to the depot vertex, the container vertices are much more complex. Though container vertices can be divided into four kinds according to the directions and states, transportation modes affect the locations of trucks so that the transportation time and time spent on the activities are different for the same vertex. The activities of these four kinds of containers are listed below in Table 1, according to which time of these activities can be calculated.

Table 1. Activities of four kinds of container vertices

Container vertex *i*	Activities
$i\in V_{IF}$	Loading the container at the **terminal/station**, delivering to the customer location, unloading and unpacking
$i\in V_{OF}$	Packing, loading the container at its customer location, delivering to the **terminal/station**, and unloading
$i\in V_{IE}$	Loading the container at the **terminal/station**
$i\in V_{OE}$	Unloading the container at the **terminal/station**

Though the time calculated above show the various kinds of time for the activities, a truck maybe spends more time for delivering a container in a vertex due to the time windows of the container and the timetables of freight trains. That is because a truck sometime has waiting time in the delivering process. This kind of time that is consumed by a container vertex is named the serving time of the vertex. The serving time is the shortest time for which a truck carries a container to finish a service. That means valueless waiting time is avoided as much as possible and the serving time cannot be compressed. The serving time of container vertices under trucking are denoted as $T^t(i)$, and the serving time of container vertices involving rail transportation are $T^s(i)$. The concepts of serving time and serving time windows below come from the paper by Zhang et al. (2009).

$$T^t(i)=\begin{cases}\max\left(\tau_{Ci}-\tau_{Bi}, tlut_i+t_{Ti}\right)+tluc_i+tpuc_i, & i\in V_{IF}\\ \max\left(\tau_{Ci}-\tau_{Bi}, tpuc_i+tluc_i+t_{iT}\right)+tlut_i, & i\in V_{OF}\\ tlut_i, & i\in V_{IE}\bigcup V_{OE}\end{cases} \tag{1}$$

$$T^s(i)=\begin{cases}\max\left(\tau_{Ci}-\tau_{Bi}-RT-tlut_i, tlus_i+t_{Si}\right)+tluc_i+tpuc_i, & i\in V_{IF}\\ \max\left(\tau_{Ci}-RT-tlus_i-\tau_{Bi}, tpuc_i+tluc_i+t_{iS}\right)+tlus_i, & i\in V_{OF}\\ tlus_i, & i\in V_{IE}\bigcup V_{OE}\end{cases} \tag{2}$$

Further, the serving time window of a container vertex is defined as the time period during which the service of this container vertex should be started. It combines the time windows of a container and the timetables of trains together. $\left[T_A^t(i), T_B^t(i)\right]$ indicates the service time window of vertex i under the situation involving only a truck, $\left[T_A^s(i), T_B^s(i)\right]$ shows the corresponding value in the other situation.

$$T_A^t(i)=\begin{cases}\min\left(\max\left(\tau_{Ai}, \tau_{Ci}-tlut_i-t_{Ti}\right), \tau_{Bi}\right), & i\in V_{IF}\\ \min\left(\max\left(\tau_{Ai}, \tau_{Ci}-tpuc_i-tluc_i-t_{iT}\right), \tau_{Bi}\right), & i\in V_{OF}\\ \tau_{Ai}, & i\in V_{IE}\\ \tau_{Ci}, & i\in V_{OE}\end{cases} \tag{3}$$

$$T_B^t(i)=\begin{cases}\min\left(\tau_{Di}-tlut_i-t_{Ti}, \tau_{Bi}\right), & i\in V_{IF}\\ \min\left(\tau_{Di}-tpuc_i-tluc_i-t_{iT}, \tau_{Bi}\right), & i\in V_{OF}\\ \tau_{Bi}, & i\in V_{IE}\\ \tau_{Di}, & i\in V_{OE}\end{cases} \tag{4}$$

$$T_A^s(i)=\begin{cases}\min\left(\max\left(\tau_{Ai}+tlut_i+RT, \tau_{Ci}-tlus_i-t_{Si}\right), \tau_{Bi}+tlut_i+RT\right), & i\in V_{IF}\\ \min\left(\max\left(\tau_{Ai}, \tau_{Ci}-tpuc_i-tluc_i-t_{iS}-tlus_i-RT\right), \tau_{Bi}\right), & i\in V_{OF}\\ \tau_{Ai}+tlut_i+RT, & i\in V_{IE}\\ \tau_{Ci}-RT-tlus_i, & i\in V_{OE}\end{cases} \tag{5}$$

$$T_B^s(i)=\begin{cases}\min\left(\tau_{Di}-tlus_i-t_{Si}, \tau_{Bi}+tlut_i+RT\right), & i\in V_{IF}\\ \min\left(\tau_{Di}-tpuc_i-tluc_i-t_{iS}-tlus_i-RT, \tau_{Bi}\right), & i\in V_{OF}\\ \tau_{Bi}+tlut_i+RT, & i\in V_{IE}\\ \tau_{Di}-RT-tlus_i, & i\in V_{OE}\end{cases} \tag{6}$$

3.1.2 Arcs

In the graphical model $G=(V,A,T,C)$, A denotes various arcs of the form (i,j) that connect vertices and the parameter T includes not only serving times of vertices but also transfer times for arcs. These arcs are decided based on the beginning vertices and ending vertices. Because the beginning vertex i and the ending vertex j have their own transportation modes, kinds of them ought to be more than vertices. The transportation modes only affect the time of these activities. Let T_{ij} denote the time for arc (i,j). These time arcs can be calculated following the Table 2 below.

Table 2. Activities of arc (i, j)

Arc (i, j)	$j \in V_D$	$j \in V_{IF} \bigcup V_{IE}$	$j \in V_{OF}$	$j \in V_{OE}$
$i \in V_D$	--	Moving **to the terminal/station**	Loading an empty container at the depot, delivering to the shipper of container c_j, and unloading	Loading an empty container at the depot, and delivering **to the terminal/station**
$i \in V_{IF}$	Loading the recently emptied container c_i, delivering to the depot ,and unloading	Loading the recently emptied container c_i, delivering to the depot, unloading and moving **to the terminal/station**	If the receiver of container c_i is not the shipper of container c_j, loading the recently emptied container c_i, delivering to the shipper of container c_j, and unloading; otherwise, doing nothing	Loading the recently emptied container c_i, and delivering **to the terminal/station**
$i \in V_{OF} \bigcup V_{OE}$	Moving **from the terminal/station** to the depot	Doing nothing	Moving **from the terminal/station** to the depot, loading an empty container, delivering to the shipper of container c_j, and unloading	Moving to the depot, loading an empty container, and delivering **to the terminal/station**
$i \in V_{IE}$	Delivering the empty container c_i **from the terminal/station** to the depot, and unloading	Delivering the empty container c_i **fro m the terminal/station** to the depot; unloading, and moving **to the terminal/station**	Delivering the empty container c_i **from the terminal/station** to the shipper of container c_j, and unloading	Doing nothing

3.1.3 Costs

In the graphical model $G = (V, A, T, C)$, the parameter *C* shows the transportation costs in the planning time-horizon. They are built up on the previous serving times and transfer times. One cost here is the sum of the cost of service of vertex *j* and the cost of transfer arc (i, j). When the service of vertex *j* involves rail, the cost should include the rail cost. Let C_{ij}^t denote the cost of arc (i, j) and vertex *j* where vertex *j* involves trucking, C_{ij}^s denote the cost of arc (i, j) and vertex *j* where vertex *j* involves rail. There should be four kinds of costs:

a) Cost of arc (i, j) where both vertex *i* and vertex *j* are served by truck;

b) Cost of arc (i, j) where vertex *i* is served by rail and vertex *j* is served by truck;

c) Cost of arc (i, j) where vertex *i* is served by truck and vertex *j* is served by rail;

d) And cost of arc (i, j) where vertex *i* is served by train and vertex *j* is served by rail.
The third case — c is used as an example here.

Table 3. The cost of arc (i, j) where vertex *i* is served by truck and vertex *j* is served by rail (c)

C_{ij}^s	$j \in V_D$	$j \in V_{IF}$	$j \in V_{IE}$	$j \in V_{OF}$	$j \in V_{OE}$
$i \in V_D$	$+\infty$	$C_t(t_{0S} + t_{Sj}) + C_s$	$C_t t_{0S} + C_s$	$C_t(t_{0j} + t_{jS}) + C_s$	$C_t t_{0S} + C_s$
$i \in V_{IF}$	$C_t t_{i0}$	$C_t(t_{i0} + t_{0S} + t_{Sj}) + C_s$	$C_t(t_{i0} + t_{0S}) + C_s$	$C_t(t_{ij} + t_{jS}) + C_s$ or $C_t t_{jS} + C_s$	$C_t t_{iS} + C_s$
$i \in V_{OF} \bigcup V_{OE}$	$C_t t_{T0}$	$C_t(t_{TS} + t_{Sj}) + C_s$	$C_t t_{TS} + C_s$	$C_t(t_{T0} + t_{0j} + t_{jS}) + C_s$	$C_t(t_{T0} + t_{0S}) + C_s$
$i \in V_{IE}$	$C_t t_{T0}$	$C_t(t_{T0} + t_{0S} + t_{Sj}) + C_s$	$C_t(t_{T0} + t_{0S}) + C_s$	$C_t(t_{Tj} + t_{jS}) + C_s$	$C_t t_{TS} + C_s$

3.2 Mathematical modeling

The mathematical programming is built up for the optimization of the problem based on the graphical model in Section 3.1.

Below are the decision variables.

$$X_{ij}^{t} = \begin{cases} 1, & \text{if arc (i,j) is included in a route and vertex j in it does not involve a train} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

$$X_{ij}^{s} = \begin{cases} 1, & \text{if arc (i,j) is included in a route and vertex j in it involves a train} \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

y_i : the start time to serve vertex i

$$\text{Objective: } \min \sum_i \sum_j \left[C_{ij}^s X_{ij}^s + C_{ij}^t X_{ij}^t \right] \quad (9)$$

$$\begin{cases}
\sum_{i \in V\setminus\{0\}} \left(X_{0j}^s + X_{0j}^t \right) = \sum_{i \in V\setminus\{0\}} \left(X_{i0}^s + X_{i0}^t \right) & (10) \\
\sum_{j \in V\setminus\{0\}} \left(X_{ij}^s + X_{ij}^t \right) = \sum_{j \in V\setminus\{0\}} \left(X_{ji}^s + X_{ij}^t \right) = 1 & (11) \\
\sum_{i,j \in Z} \left(X_{ij}^s + X_{ij}^t \right) \le |Z| - 1 \qquad \forall Z \subseteq V\setminus\{0\}, Z \ne \Phi & (12) \\
\left(\sum_{i \in V-Z} \sum_{j \in V} \left(X_{ij}^s + X_{ij}^t \right) \ge 1 \qquad \forall Z \subseteq V\setminus\{0\}, Z \ne \Phi \right) & (13) \\
T_A^s(j) - \left(1 - X_{ij}^s\right) M \le X_{ij}^s y_j \le T_B^s(j) + \left(1 - X_{ij}^s\right) M \qquad \forall i, j \in V\setminus\{0\} & (14) \\
T_A^t(j) - \left(1 - X_{ij}^t\right) M \le X_{ij}^t y_j \le T_B^t(j) + \left(1 - X_{ij}^t\right) M \qquad \forall i, j \in V\setminus\{0\} & (15) \\
X_{ij}^s \left(y_j - y_i - T^s(i) - T_{ijs} \right) + X_{ij}^t \left(y_j - y_i - T^t(i) - T_{ijt} \right) \ge 0 \qquad \forall i, j \in V\setminus\{0\} & (16) \\
X_{ij}^s \in \{0,1\}, X_{ij}^t \in \{0,1\}, \ y_i : \text{ real variable}, \qquad \forall (i,j) \in A & (17)
\end{cases}$$

The objective function (9) minimizes the total transportation cost. The constraint (10) indicates that the number of trucks that starts from the depot should be exactly the same as the number returning to the depot.

The constraint (11) means every vertex (except the depot) should be served exactly once.

The constraint (12) eliminates the sub-tours among container vertices. Φ refers to empty set.

The constraint (13) implies that for any subset $\forall Z \subseteq V\setminus\{0\}$, there is at least one passage connecting subset Z.

Constraints (11) and (12) together have the same effect as constraints (11) and (13). Thus, one of (12) and (13) can be chosen.

Constraints (14) and (15) mean that the service of any container vertex should be started during its corresponding time window. $\left[T_A^s(i), T_B^s(i)\right]$ indicates the service time window of vertex *i* under the situation involving a train, $\left[T_A^t(i), T_B^t(i)\right]$ shows the other situation.

The constraint (16) updates the start time along the route.

4. Solutions

A hybrid SA algorithm is coded in MATLAB 7.0. The basic principle of SA comes from Baker (2001). The performance of SA is often hindered by its slow convergence to optimal or near-optimal solution (Sait and Youssef, 1999). Therefore, a greedy algorithm is used to obtain the initial solution.

Let $T(i)$ denote the temperature at stage *i* and Z_i denote the total transportation cost of scheduled workload sequence at stage *i*. $T(i+1) = T(i) \times \pi$. The parameter π is used to control the temperature to go down step by step. Its value is between 0 and 1. If the temperature of the process gets lower than the ending temperature, the process stops.

The probability that the *j*th neighbor at stage *i* becomes the next seed is $q_{ij} = \min\left\{1, e^{-\Delta Z/T(i)}\right\}$, where $\Delta Z = Z_j - Z_i$. Figure 1 shows the logic of hybrid SA in this paper.

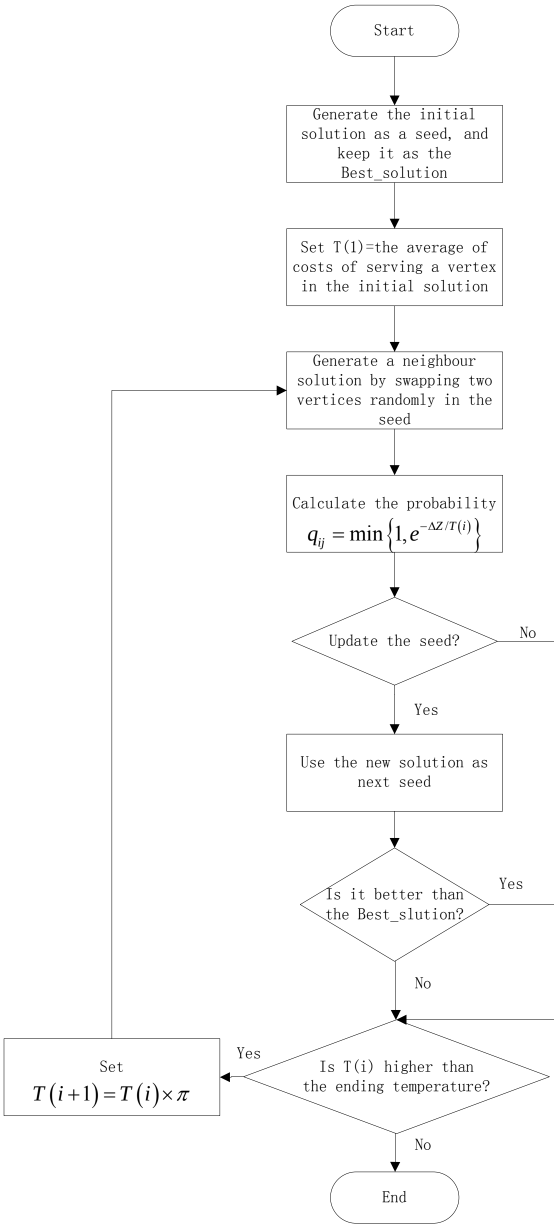


Figure 1. Procedure of the hybrid SA algorithm

An initial solution is obtained as follows.

Step 1: Build up four time-tables and four cost tables above, in which the first vertex is the depot vertex.

Step 2: Set the current vertex *i* as the depot.

Step 3: Select the next vertex *j*.

Step 3.1: Update the elapsed time.

Step 3.2: Select candidates from the remaining vertices whose serving time-windows can be satisfied.

Step 3.3: Choose those with the lowest cost from each cost table. Compare them and continue to choose one with the lowest cost as the next vertex *j*. If there is more than one vertex with the lowest cost, the smallest vertex index is selected as the next vertex *j*.

Step 4: If *j* is the depot, go to step 5; otherwise mark *j* as a scheduled vertex, set the current vertex *i*=*j*, and go to step 3.

Step 5: If all vertices are scheduled, stop; otherwise, go to Step 2.

5 Numerical experiments

In this section, numerical examples are offered and the results are compared and analyzed.

5.1 Example setting

The data used in this paper are generated by the similar way in Zhang et al. (2009).

1) Types of containers are generated randomly and each type has the same probability.
2) The loading/unloading times for an empty container at the depot, the terminal, and the customer locations are distributed within the range of 2 min to 5 min.
3) The packing/unpacking times for a full container are distributed within the range of 5 min and 60 min. For an empty container, this time is 0.
4) The traveling time between any two locations is generated in the interval from 0 to 6 h.
5) The lower bounds of the pick-up time windows of all the containers are distributed within the range of 8:00 am to 12:00 am. The interval lengths of pick-up time windows are generated as random variables in the interval from 0 to 3 h. The lower bounds of the delivery time windows for full containers are calculated by the corresponding lower bounds of the pick-up time windows plus the traveling times between the corresponding customer locations and the terminal. The interval lengths of the delivery time windows are distributed within the range of 2 h and 5 h. Inbound empty containers only have pick-up time windows and outbound empty containers have only delivery time windows.
6) The times of trains are generated randomly between 2 h and 6 h when vertex information is generated.
7) The cost for each container under rail is calculated by multiplying the train-travel time by a coefficient 0.3.
8) Time data are converted into integers from the beginning of a day. For example, 8:00 is converted into 480.

The serving times and serving time windows are calculated and the transfer time tables and cost time tables can be calculated based on the above data.

An instance including five container vertices is randomly generated and cited in Table 4. The corresponding traveling times are given in Table 5.

Table 4. Time windows of an instance including 5 containers

NO	Type	Customer ID	TW			
			Pl	Pu	Dl	Du
1	IF	3	710	841	941	941
2	OE	0			747	753
3	OF	2	530	639	762	686
4	OF	1	487	543	642	510
5	OF	5	627	737	737	913

(OE—Outbound empty container, IE—Inbound empty container, OF—Outbound full container, IF—Inbound full container; Customer ID=0 means depot because requests regarding IE and OE are made by the depot.)

Table 5. Traveling times between pairs of vertices (D—Depot, T—Terminal, S—Station)

	D	T	S	1	2	3	4	5
D	0	49	169	282	0	191	120	21
T	49	0	4	231	49	156	23	286
S	169	4	0	81	169	356	322	166
1	282	231	81	0	282	107	210	186
2	0	49	169	282	0	191	120	21
3	191	156	356	107	191	0	71	274
4	120	23	322	210	120	71	0	137
5	21	286	166	186	21	274	137	0

(It is assumed that locations of OE and IE are at the depot.)

The travel time by train is 225 min and the cost by train is 67.5 per container. The cost per minute under trucking is 1.

5.2 Experimental results

The experiments are tested on a computer with Intel® Core™2 Quad CPU Q8400 @ 2.66GHz 2.67GHz, 2.00GB memory. In the hybrid SA algorithm, the ending time is set to 1 and $\pi = 0.9$.

Table 6. Comparison of the three solution methods

Example NO	Amount of containers	Amount of Customers	Hybrid SA				
			Obj. value	CPU time (s)	Amount of trucks	Served by truck	Served by rail
1	5	5	1096.5	30.656	2	4	1
2	10	10	4048.2	45.337	5	6	4
3	20	20	7305.8	92.563	8	18	2
4	30	20	13335	106.859	15	19	11
5	40	20	11975	276.672	18	33	7
6	50	30	14392	715.625	26	40	10
7	60	40	21854	1341.516	29	30	30
8	70	50	31336	2660.828	40	53	17
9	80	60	23841	4375.453	41	71	9

The column "Amount of trucks" means the amount of trucks used in the schedule. The column "Served by rail" represents the amount of containers involving train transportation while "Served by truck" represents the others only involving trucks.

The hybrid SA algorithm takes longer and longer time to get an acceptable solution as the problem becomes bigger and bigger. One of the reasons is that it needs more time to obtain an initial solution because containers are assigned to trucks one by one under the time window constraints. For the very large problem, maybe we can reduce the time by dividing it into several sub problems and combining them at last. That means we sacrifice some efficiency for shorter time.

6. Conclusions

A container transportation problem under trucking and rail transportation modes has been studied. One depot, one station, and one terminal are considered and four types of containers are involved with time windows. A graphical model was developed to formulate a mathematical programming. A hybrid SA algorithm was proposed to solve this problem and numerical examples are also studied to investigate the performance of the proposed algorithm. For further studies, we should compare the algorithm with other heuristics.

Acknowledgements

"This work was supported by a Grant from the Korean Ministry of Education, Science and Technology" (The Regional Core Research Program/Institute of Logistics Information Technology).

7. References

Baker Kenneth R. (2001). Elements of sequencing and scheduling: 4.11-4.14

Jula H., Dessouky M., Ioannou P. and Chassiakos A. (2005). Container movement by trucks in metropolitan networks: modeling and optimization. *Transportation Research*, Vol. 41: 235-259

Liu X. and He G.G. (2007). A research on tabu search for vehicle routing problem. *Computer Engineering and Applications*, Vol. 24: 179-181 and 199

Sait Sadiq M. and Youssef Habib. (1999). Iterative computer algorithms with applications in engineering—solving combinatorial optimization problems. *IEEE Computer Society Press*, 349

Wen S, Zhou P. (2007). A container vehicle routing model with variable traveling time. *IEEE International Conference on Automation and Logistics:* 2243-2247

Zhang R., Yun W.Y. and Moon I.K. (2009). A reactive tabu search algorithm for the muti-depot container truck transportation problem. *Transportation Research Part E: Logistics and Transportation Review*, Vol.45: 904-914

HIERARCHICAL APPROACH TO VEHICLE ROUTING PROBLEM FOR DELIVERY AND INSTALLATION

Kyoung Cheol Kim[1] and Shiwoo Lee [2]*

[1]School of Mechanical, Industrial & Manufacturing Engineering,
Oregon State University,
Corvallis, OR 97331, USA
e-mail: kimkyo@engr.oregonstate.edu

[2]Division of Mechanical and Industrial Engineering
Pohang University of Science and Technology
Pohang, Korea
e-mail: cameolee@postech.ac.kr

Abstract: The vehicle routing and scheduling problems have been studied extensively for various industries with special needs. In this paper, a vehicle routing problems considering unique characteristics of electronics industry is considered. A mixed-integer nonlinear programming (MINP) model has been presented to minimize the traveling time of delivery and installation vehicles. A hierarchical approach using the genetic algorithm has been proposed and implemented to solve problems of various sizes. The computational results show the effectiveness and the efficiency of the proposed hierarchical approach. A performance comparison between the MINP approach and the hierarchical approach is presented.

1. INTRODUCTION

In order to survive in this competitive business environment, a company must have ways to handle various materials of concern cost-effectively. Especially in manufacturing industries, material handling methods for raw materials and works-in-process are as important as the ones for final products. In order that material handling activities satisfy business' and customers' demands effectively, the vehicle routing and scheduling problems have been studied and implemented extensively for various industries with special needs (Golden and Wasil, 1987; List and Mirchandani, 1991; Spasovic *et al*, 2001; Zografos and Androutsopoulos, 2004; Ripplinger, 2005; Prive *et al*, 2006; Claassen and Hendricks, 2007; Ji, 2007). In this study, a variant of the vehicle routing problem (VRP), which has been characterized in the electronics industry to satisfy its unique material handling needs as the paradigm of distribution has been shifted from the past, has been presented.

In recent days, the electronics industry experiences rapidly-emerging changes in their post-sales service, i.e., delivery and installation. These trends tend to add the responsibilities of delivery and installation onto electronics manufacturers, and the number of direct deliveries from electronics manufacturers to customers increases at an explosive pace.

In addition, many newly-emerging products require not only the delivery, but also the professional installation; these products include wall-mounting televisions (home theater systems), washers and dryers, refrigerators with water purifying systems, special cook-tops, numerical control machines, computer servers, and etc.

There has been another organizational need that makes the task of planning for vehicle routing and scheduling more complicated. The expense to maintain manufacturers' nation-wide distribution and service network over customers is too high to make it practical and economical. Therefore, manufacturers adopt the practice of outsourcing the delivery to third parties and maintaining their own service teams or commissioning authorized service providers for the installation service.

The VRP under consideration in this paper assumes that there exist two types of demands in a complex electronics market: one type requires only the delivery, and the other requires both the delivery and the installation. Because it is not efficient to have the skilled professionals with installation capability to drive the delivery vehicles, haul the (potentially heavy) products, and provide the installation service, electronics manufacturers prefer to operate separately two different types of vehicles (delivery and installation vehicles). It is assumed that delivery vehicles have a limited loading capacity to carry the products and installation vehicles do not. Both types of vehicles start from a depot at the beginning and return to the depot within a specified duration of time, i.e., *maximum operation time*. All delivery demands of customers are known in advance. The sum of demands of customers assigned to a delivery vehicle cannot exceed the loading capacity of the delivery vehicle. For the customers requiring only the delivery, a delivery vehicle is allowed to visit them only once. For the customers requiring the delivery and the installation, a delivery vehicle and an installation vehicle are allowed to visit them only once, respectively. In addition, there is an important constraint to guarantee the quality of service, i.e., *service level*, which is defined as a maximum amount of time lapse between the arrivals of delivery and installation vehicles. An installation vehicle must visit a customer within the service level after the visit (or arrival) of a delivery vehicle to (at) that customer. Therefore, the synchronization of both types of vehicles is required to ensure the guaranteed quality of service for customers ordering both the delivery and the installation. Figure 1 shows an example of the VRP under consideration and its potential solution. The example consists of 16 customers; 10 customers requiring the delivery-only (which are represented by the filled circles) and 6 customers requiring both delivery and installation (which are represented by the clear circles). The solid and dotted lines are the routes for delivery and installation vehicles, respectively. There are three delivery and two installation vehicles in Figure 1. The arrival times of delivery and installation vehicles are shown next to the customers. The installation vehicles can visit customers earlier than the delivery vehicle, which produces waiting times for installation vehicles at the corresponding customer locations. If the installation vehicles visit customers later than the delivery vehicles, the service level must be satisfied.

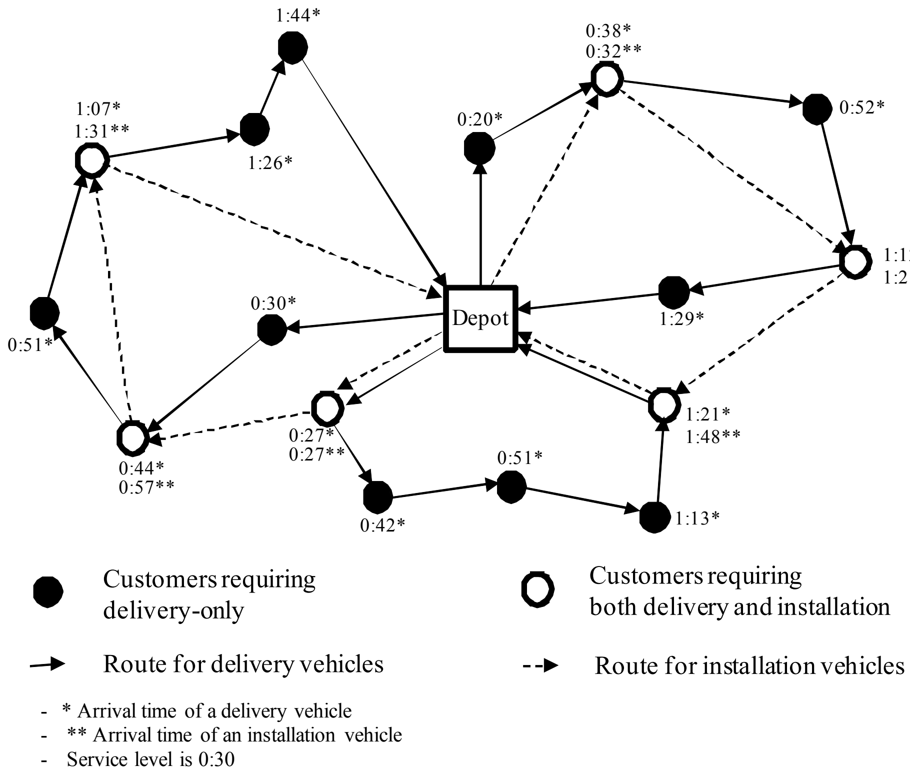


Figure 1. An example of the vehicle routing problem under consideration

The VRP under consideration can be viewed as a combination of two traditional VRPs; one is the VRP for delivery vehicles, which has the characteristics of capacitated VRP (Laporte *et al*, 1987; Campos and Mota, 2000; Toth and Vigo, 2002; Pisinger and Ropke, 2007), and the other is the VRP for installation vehicles, which has the characteristics of VRP with time windows (Kolen *et al*, 1987; Potvin and Rousseau, 1995; Bräysy and Gendreau, 2005).

A mathematical programming model for this problem has been formulated and solved using commercially available optimization software. The computational results show that it is difficult to obtain the optimal solutions for even problems of small size. A hierarchical approach using the genetic algorithm (GA) has been proposed to solve the problem under consideration efficiently in a reasonable amount of time. The hierarchical approach divides the problem into two subproblems: the VRP for delivery vehicles in Stage 1 and the VRP for installation vehicles in Stage 2. The vehicle routes and schedules for delivery vehicles are determined in Stage 1. Based on the results of Stage 1, the vehicle routes and schedules for installation vehicles are obtained in Stage 2. The computational results also show that the best results of Stage 1 do not always lead to the best results for the original problem. Various examples have been tested to demonstrate the effectiveness of the proposed hierarchical approach and the computational results are given in Section 4.

This study can provide a management strategy to efficiently plan the vehicle routes and schedules for delivery and installation vehicles and to effectively synchronize those vehicles.

2. MATHEMATICAL PROGRAMMING MODEL

In this section, a mathematical programming model of the VRP under consideration is presented. The objective of this model is to minimize the shortest distance or time of travels by all delivery and installation vehicles. The decision variables provide optimal routes and schedules of delivery and installation vehicles. The VRP under consideration includes a depot (at location 0) and N customers (at locations 1 to N), among which customers belong to A require both the delivery and the installation ($|A| \le N$).

The notations for decision variables and parameters are explained in the following:

Decision Variables:

x_{ijp} $\begin{cases} 1, & \text{if the delivery vehicle p travels from location i to location j} \\ 0, & \text{otherwise} \end{cases}$

y_{ijq} $\begin{cases} 1, & \text{if the installation vehicle q travels from location i to location j} \\ 0, & \text{otherwise} \end{cases}$

e_i Arrival time of the delivery vehicle at location i, $i \in A$

f_i Arrival time of the installation vehicle at location i, $i \in A$

w_i Waiting time of the installation vehicle at location i, $i \in A$

u_{ip} Subtour prevention variables for x_{ijp}

v_{iq} Subtour prevention variables for y_{ijq}

Parameters:

N	Number of customers
K	Number of delivery vehicles
S	Number of installation vehicles
A	Set of locations of customers requiring both delivery and installation
T_{ij}	Traveling time from location i to location j
D_i	Demand of customer at location i
VC	Capacity of a delivery vehicle
R_i	Time to complete an installation at location i, $i \in A$
OT	Available operation time per shift for vehicles
SL	Service level

The problem under consideration is formulated as a mixed-integer nonlinear programming (MINP) model as follows:

$$\text{Minimize } Z = \sum_{p=1}^{K}\sum_{i=0}^{N}\sum_{j=0}^{N} T_{ij} x_{ijp} + \sum_{q=1}^{S}\sum_{i=0}^{N}\sum_{j=0}^{N} T_{ij} y_{ijq}$$

Subject to

$$\sum_{p=1}^{K}\sum_{j=1}^{N} x_{ijp} \le K \quad \text{for } i = 0 \qquad (1)$$

$$\sum_{j=1}^{N} x_{ijp} \le 1 \quad \text{for } i = 0, \forall p \qquad (2)$$

$$\sum_{j=1}^{N} x_{ijp} - \sum_{j=1}^{N} x_{jip} = 0 \quad \text{for } i = 0, \forall p \qquad (3)$$

$$\sum_{p=1}^{K}\sum_{j=0}^{N} x_{ijp} = 1 \quad \text{for } i = 1 \ldots N \qquad (4)$$

$$\sum_{p=1}^{K}\sum_{i=0}^{N} x_{ijp} = 1 \quad \text{for } j = 1 \ldots N \qquad (5)$$

$$\sum_{j=0}^{N} x_{ijp} - \sum_{j=0}^{N} x_{jip} = 0 \quad \text{for } i = 1 \ldots N, \ \forall p \qquad (6)$$

$$\sum_{i=1}^{N} D_i \left(\sum_{j=0}^{N} x_{ijp} \right) \le VC \quad \text{for } \forall p \qquad (7)$$

$$\sum_{i=0}^{N}\sum_{j=0}^{N} T_{ij} x_{ijp} \le OT \quad \text{for } \forall p \qquad (8)$$

$$u_{ip} - u_{jp} + (N+1)x_{ijp} \le N \quad \text{for } i \ne 0, j \ne 0, i \ne j, \forall p \qquad (9)$$

$$\sum_{i=1}^{N} x_{iip} = 0 \quad \text{for } \forall i, \forall p \qquad (10)$$

$$x_{ijp} = \{0,1\} \quad \text{for } \forall i, \forall j, \forall p \qquad (11)$$

$$\sum_{q=1}^{S}\sum_{j=1}^{N} y_{ijq} \le S \quad \text{for } i = 0, j \in A \qquad (12)$$

$$\sum_{j=1}^{N} y_{ijq} \le 1 \quad \text{for } i = 0, \forall q \qquad (13)$$

$$\sum_{j=1}^{N} y_{ijq} - \sum_{j=1}^{N} y_{jiq} = 0 \quad \text{for } i = 0, \forall q \qquad (14)$$

$$\sum_{q=1}^{S}\sum_{j=0}^{N} y_{ijq} = 1 \quad \text{for } i \in A \qquad (15)$$

$$\sum_{q=1}^{S}\sum_{i=0}^{N} y_{ijq} = 1 \quad \text{for } j \in A \qquad (16)$$

$$\sum_{j=0}^{N} y_{ijq} - \sum_{j=0}^{N} y_{jiq} = 0 \quad \text{for } i \in A, \forall q \qquad (17)$$

$$v_{iq} - v_{jq} + (N+1)y_{ijq} \le N \quad \text{for } i \ne 0, j \ne 0, i \ne j, \forall q \qquad (18)$$

$$\sum_{i=1}^{N} y_{iiq} = 0 \quad \text{for } i \in A, \forall q \qquad (19)$$

$$y_{ijq} = \{0,1\} \quad \text{for } \forall i, \forall j, \forall q \qquad (20)$$

$$e_0 = f_0 = w_0 = R_0 = 0 \qquad (21)$$

$$f_i - e_i \le SL \quad \text{for } i \in A \qquad (22)$$

$$w_i \ge 0, \; e_i - f_i - w_i \le 0 \quad \text{for } i \in A \qquad (23)$$

$$\sum_{p=1}^{K}\sum_{i=0}^{N} x_{ijp}\left(e_i + T_{ij}\right) - e_j = 0 \quad \text{for } \forall j \qquad (24)$$

$$\sum_{q=1}^{S}\sum_{i=0}^{N} y_{ijq}\left(f_i + w_i + R_i + T_{ij}\right) - f_j = 0 \quad \text{for } j \in A \qquad (25)$$

$$\sum_{i=0}^{N}\sum_{j=0}^{N} y_{ijq}\left(T_{ij} + w_i + R_i\right) = OT \quad \text{for } \forall q \qquad (26)$$

The objective function is to minimize the traveling times of delivery and installation vehicles. The traveling distance or the transportation cost can be used in the objective function instead of traveling times. The constraints can be classified into three different sets; the first concerning the VRP for delivery vehicles (constraints (1)~(11)); the second concerning the VRP for installation vehicles (constraints (12)~(20)); and the third concerning the synchronization for both types of vehicles (constraints (21)~(26)). Constraints (21)-(25) guarantee the quality of service by satisfying the service level (SL) for customers requiring both the delivery and the installation.

If an installation vehicle gets to a location earlier than a delivery vehicle, it sometimes needs to wait for the delivery vehicles before the installation starts. An installation vehicle may not leave immediately after the installation at a location to avoid waiting at the next location or to make other arrangements. The waiting time (w_i) of the installation vehicle at a location is defined as an amount of time spent before or after the installation at the corresponding location. We also tested another mathematical programming model by replacing constraint (23) with another constraint, $\max\{0, e_i - f_i\} - w_i = 0$, in order to remove the waiting time after the installation at a location. Since the third set of constraints includes nonlinear ones, the mathematical programming for the VRP under consideration is more complicated than the original VRP that can be formulated as a mixed integer programming (MIP) model.

3. HIERARCHICAL APPROACH USING THE GENETIC ALGORITHM FOR SYNCHRONIZATION OF DELIVERY AND INSTALLATION

Therefore, it is necessary to develop a systematic approach not only to find routes and schedules for delivery and installation vehicles but also to synchronize both types of vehicles. A search of routes and schedules for delivery vehicles and the other search of routes and schedules for installation vehicles can be defined, as subproblems of the original problem, respectively. In order to obtain good solutions of the original problem, a hierarchical approach dealing with two subproblems is presented in this paper.

The proposed hierarchical approach divides the original problem into two stages and each stage contains a subproblem. The subproblem in Stage 1 is a VRP for delivery vehicles, and the other subproblem in Stage 2 is a VRP for installation vehicles. From the subproblem in Stage 1, a set of routes and schedules for the delivery vehicles is generated. The generated set of routes and schedules is a partial solution of the original problem and is also used as input data of the subproblem in Stage 2. Based on the partial solution from the subproblem in Stage 1, a set of routes and schedules for the installation vehicles is determined in the problem in Stage 2. The set of routes and schedules for installation vehicles is the other partial solution of the original problem. The synchronization of two types of vehicles is automatically completed while solving the subproblem in Stage 2. Finally, two partial solutions from the subproblems are consisted of the solution to the original problem.

The subproblem in Stage 1 involves the characteristics of capacitated vehicle routing problems (CVRP). The subproblem assumes that all delivery vehicles have an identical loading capacity. They must return to the depot within a specific time, called *maximum operation time*. The limited loading capacity and the maximum operation time need to be considered when customers are assigned to delivery vehicles. All customers can be visited only once by a single delivery vehicle. It also assumes that *unloading times* for all customers do not exist in this paper but it can be modified without loss of generality. *Fixed cost per delivery vehicle* is considered in the proposed algorithm to minimize the number of delivery vehicles in operation, but it can be omitted. The algorithm for the subproblem in Stage 1 determines routes and schedules for the delivery vehicles and their arrival time at all customers. The arrival times of delivery vehicles at the customers who require both the delivery and the installation are later fed to the subproblem in Stage 2, in order to be considered for the synchronization of delivery and installation vehicles.

The subproblem in Stage 2 includes the characteristics of vehicle routing problems with time windows (VRPTW). The loading capacity of installation vehicles is not considered since the installation requires the service to be rendered, not the goods. Similar to the subproblem in Stage 1, all installation vehicles must return to the depot within the specified maximum operation time. Installation vehicles must visit all customers requiring both the delivery and the installation. The customers must be visited only once by a single installation vehicle within a specified lapse after a delivery vehicle arrives at the corresponding customers. That is, installation vehicles must visit customers for the installation service within a specified duration after the delivery, so as to guarantee the quality of service. This specified duration is defined as *service level*. Hence, each customer requiring both delivery and installation has a time window for which he/she expects a visit (or arrival) by an installation vehicle. It assumes that the lengths of the time windows for all customers are identical but it can be relaxed without loss of generality. If an installation vehicle arrives earlier than a delivery vehicle for the customer or arrives earlier than the beginning of the time window for the corresponding customer, the installation vehicle must wait at the location of the customer before it starts the installation service. The waiting times of installation vehicles can be considered as a penalty. It is assumed that installation service requires the same amount of time, *installation time,* for each customer. *Fixed cost per installation vehicle* is considered in the proposed algorithm to minimize the number of installation vehicles used. The algorithm for the subproblem in Stage 2 determines routes and schedules for all installation vehicles.

Finally, for a solution of the original problem, a set of routes and schedules for all delivery and installation vehicles are decided through this hierarchical approach using genetic algorithms. The fitness function in the GA for Stage 1 considers the travel times and the fixed costs of all delivery vehicles while the one for Stage 2 considers the travel times, waiting times and the fixed costs of all installation vehicles.

The GA has been applied to many combinatorial optimization problems successfully in the past (Reeves, 1994; Jung and Haghani, 2000; Ozdemir and Mohan, 2004; Marian *et al*, 2006). The GA does not guarantee the optimality due to its stochastic nature, but it finds good near-optimal solutions in significantly lesser computational time. To effectively obtain good near-optimal solutions for the problem under consideration, the GA has been used in the hierarchical approach. For each subproblem, a one-dimensional array has been used to represent a potential partial solution. Figure 2 illustrates the genetic representation of an individual in the population for the subproblem in Stage 1.

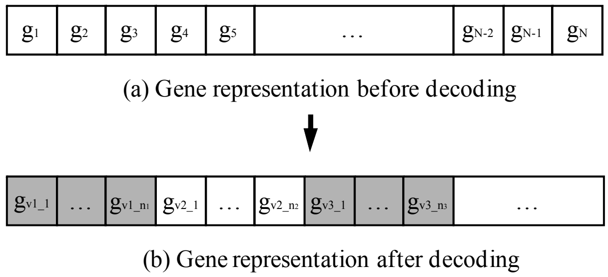


Figure 2. The genetic representation for the subproblem in Stage 1

Let N be the number of customers for the problem under consideration and n_i be the number of customers who will be assigned and served by delivery vehicle i. An individual consists of N genes (g_1, g_2, …, g_N) in Figure 2(a). The genes contain the indices of customers requiring the delivery. Since each customer is allowed to be visited by a single delivery vehicle, the index of a customer must be shown only once in the genetic representation. The decoding procedure in the GA for the subproblem in Stage 1 determines the route and schedules of each delivery vehicle in operation. In order to assign customers to a delivery vehicle, a greedy method has been used in consideration of the customers' demands, the loading capacity of delivery vehicles, and the maximum operation time. After the decoding process, the individual is interpreted as Figure 2(b). A set of customers, who are indicated by genes from g_{v1_1} to $g_{v1\_n_1}$, are allocated to the first delivery vehicle (v_1). The remaining customers are assigned to other delivery vehicles in the same manner. The alternating shade in Figure 2(b) represents groups of customers served by different delivery vehicles. The genetic representation of an individual in the population for the subproblem in Stage 2 has been developed in a similar manner and omitted due to the redundancy.

In the subproblem in Stage 1, the traveling times and the fixed costs of used delivery vehicles are considered to calculate a fitness value of an individual. Let τ_a be the sum of the traveling times of all delivery vehicles and δ_a be the sum of the fixed costs of delivery vehicles used in individual *a*. The fitness value of individual *a* for the subproblem in Stage 1 (π_a) is defined as follows;

$$\pi_a = \frac{1}{\tau_a + \delta_a}.$$

In the subproblem in Stage 2, the sum of the traveling times, the sum of the waiting times and the sum of the fixed costs of installation vehicles in operation, are considered to calculate a fitness value of an individual. Let φ_b be the sum of the traveling times of installation vehicles, ω_b be the sum of the waiting times of installation vehicles, and γ_b be the sum of the fixed costs of used installation vehicles in individual *b*. The fitness value of individual *b* for the subproblem in Stage 2 (ρ_b) is defined as follows;

$$\rho_b = \frac{1}{\varphi_b + \omega_b + \gamma_b}.$$

An individual which has a higher fitness value than others has a higher chance to participate in reproduction. The roulette-wheel selection procedure is used in the proposed GA. The probability for an individual to be selected for reproduction is the fitness value of the individual over the sum of all individuals' fitness values in the population.

A hybrid crossover procedure has been proposed and implemented to efficiently and effectively reproduce new offspring from two parents in the current population.

The proposed hybrid crossover procedure is described as follows.

Step 1: Two individuals (P1 and P2) in the current population are randomly selected.

Step 2: A vehicle used in one parent (P1) is randomly chosen. The sequence of customers' indices (or genes) assigned to the corresponding vehicle are copied to the beginning of the new offspring. The copied genes will be deleted from P1 and P2.

Step 3: Perform the one-point crossover over the remaining genes of P1 and P2.

Step 4: Append the result of Step 3 to the new offspring from Step 2.

An exchange mutation procedure has been used in the proposed algorithm by swapping two genes from an individual selected randomly at the mutation rate (Gen and Cheng, 2000). The ranking replacement strategy has been used to construct the new population for the next generation in the proposed algorithm (Chu and Beasley, 1998).

The proposed GA terminates when the number of generations reach a specified limit or no improvement of the best solution is observed over a specified number of generations, which is called *improvement interval*. The individual which has the highest fitness in the final generation is interpreted as the best known solution for the problem.

4. COMPUTATIONAL RESULTS AND CONCLUSION

All computational experiments are carried out on a personal computer with a 3.4 GHz Pentium 4 CPU and 2.0 GB RAM. The MINP model in Section 2 has been programmed and solved by Lingo version 10.0, one of commercially available software for linear and non-linear programming models. The proposed hierarchical approach using the genetic algorithm in Section 3 has been implemented in C++ programming language using the Microsoft Visual Studio.NET Framework 1.1 version.

All customers are randomly located in a 100x100 square area and a single depot is located in the center. The hierarchical approach using the genetic algorithm has solved the problem through two stages. For the GA used in the proposed hierarchical approach, the following parameters are used. The size of the population is 200. The crossover rate and the mutation rate are set to 0.8 and 0.1, respectively. The GA terminates if the number of generation reaches 1000 or there is no better solution than the current best solution in the improvement interval of 200.

An exemplary test problem has been solved and presented to show the effectiveness of the hierarchical approach using the genetic algorithm. There are 30 customers requiring the delivery while 10 customers require the installation. The delivery demand of each customer is between 2 and 10; there are 8 delivery vehicles at the depot and the loading capacity of delivery vehicles is 20; there are 3 installation vehicles at the depot; the installation time and the service level are 10 and 60, respectively; and the maximum operation time is 480.

Figure 3 shows the best solution produced by the proposed hierarchical approach. The solution uses 8 delivery vehicles and 2 installation vehicles. The proposed algorithm has been completed in 1.87 seconds and the sum of the travelling time of delivery and installation vehicles is 1111.43. The routes from and to the depot have been omitted for the simplicity in Figure 3.

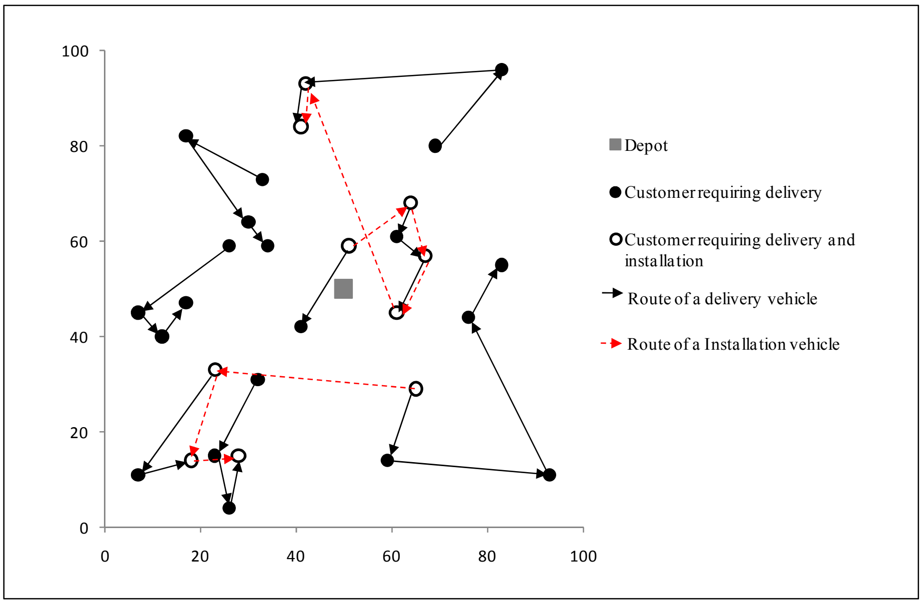


Figure 3. The routes of delivery and installation vehicles in the best solution of example 1

In this paper, a new type of VRPs in electronics industries has been identified and introduced. The problem under consideration has two types of demands: one type requires only the delivery, and the other requires both the delivery and the installation. To satisfy both demands, electronics manufacturers prefer to operate separately two different types of vehicles (delivery and installation vehicles). There is an important constraint to guarantee the quality of service, that is, an installation vehicle must visit a customer within a specific duration after the visit (or arrival) of a delivery vehicle to (at) that customer. Therefore, the synchronization of both types of vehicles is required to ensure the guaranteed quality of service for customers ordering both the delivery and the installation.

The MINP model for the problem under consideration has been formulated and solved using the commercially available software, Lingo. In addition, a hierarchical approach using the genetic algorithm has been proposed and developed to solve problems of significant sizes. The MINP model using Lingo has only been able to find the optimal solutions for small problems in a considerable amount of time. To verify the effectiveness and performance of the proposed hierarchical approach using the genetic algorithm, various test problems have been randomly created. The computational results show that the proposed hierarchical approach is able to find the optimal or near-optimal solutions to problems of various sizes in a reasonable amount of time.

In addition, the computational results show that the result of the subproblem in Stage 2 is subject to the result of the subproblem in Stage 1 in the proposed hierarchical approach. The best solution of the subproblem in Stage 1 does not guarantee that it produces the best solution of the original solution. The future research will be performed in a direction to considering both subproblems at the same time.

5. REFERENCES

Bräysy O, and Gendreau M (2005). Vehicle routing problem with time windows, Part I: Route construction and local search algorithms. *Transport Sci* **39:** 104-118.

Campos V and Mota E (2000). Heuristic procedures for the Capacitated Vehicle Routing Problem. *Comput Optim Appl* **16:** 265-277.

Chu P C and Beasley J E (1998). Constraint handling in genetic algorithms: the set partitioning problem. *J Heuristics* **4:** 323-357.

Gen M and Cheng R (2000). *Genetic algorithms and engineering optimization*. John Wiley & Sons, Inc.

Golden B L and Wasil E A (1987). Computerized Vehicle Routing in the Soft Drink Industry. *Opns Res* **35:** 6-17.

Kolen A, Rinnooy Kan A and Trienekens H (1987). Vehicle routing with time windows. *Opns Res* **35:** 266-273.

Laporte G, Nobert Y and Desrochers M (1985). Optimal routing under capacity and distance restrictions. *Opns Res* **33:** 1050-1073.

List G and Mirchandani P (1991). An Integrated Network/Planner Multiobjective Model for Routing and Scheduling for Hazardous Materials and Wastes. *Transport Sci* **25:** 146-156.

Marian R, Luong L and Abhary K (2006). A genetic algorithm for the optimisation of assembly sequences. *Comput Indust Eng* **50:** 503-527.

Ozdemir H T and Mohan C K (2001). Flight graph based genetic algorithm for crew scheduling in airlines. *Inf Sci* **133:** 165-173.

Pisinger D and Ropke S (2007). A general heuristic for vehicle routing problems. *Comput Opns Res* **34:** 2403-2435.

Potvin J Y and Rousseau J M (1995). Exchange heuristic for routeing problems with time windows. *J Opl Res Soc* **46:** 1433-1446

Prive J, Renaud J and Boctor F (2006). Solving a vehicle-routing problem arising in soft-drink distribution. *J Opl Res Soc* **57:** 1045-1052.

Reeves C R. (1994). Genetic algorithm for flowshop sequencing. *Comput Opns Res* **22:** 5-13.

Ripplinger D (2005). Rural school vehicle routing problem *Transport Res Rec* **1922:** 105-110.

Toth P and Vigo D (2002). Models, relaxations and exact approaches for the capacitated vehicle routing problem. *Discrete Appl Math* **123:** 487-512.

Zografos K G and Androutsopoulos K N (2004). A heuristic algorithm for solving hazardous materials distribution problems. *Eur J Opl Res* **152:** 507-519.

Opening Ceremony

·**Day1: Sep. 15 (Wed.)**

·**Time: 10:30 - 12:20**

·**Chair: Kap Hwan Kim**

·**Room: Grand Ballroom, 5F**

FUTURE FOR THE AUTOMATION OF CONTAINER TERMINALS

Michael G H Bell
Professor of Transport Operations
Port Operations Research and Technology Centre
Department of Civil Engineering
Imperial College London
m.g.h.bell@ic.ac.uk

Abstract: Container terminal automation is gradually spreading, starting in Europe with Rotterdam and Hamburg. This presentation looks at the evolving technology for automation with respect to the different modes of container terminal operation. For straddle carrier terminals, it looks at the Autostrad from Patrick, its performance so far in Brisbane, and prospects for adoption elsewhere. For rail mounted gantry terminals, it looks at the popularity of the Automatic Stacking Crane and the different configurations to be found. Different forms of automated horizontal transport, ranging from Automatic Guided Vehicles to Automatic Lifting Vehicles, are presented along with their advantages and disadvantages. It is noted that Automatic Lifting Vehicles effectively decouple horizontal from vertical transport, allowing a smaller fleet size, however the technology is more complex, costly and less reliable. The semi-automation of quay cranes is also described. The economic and environmental drivers behind automation are analysed, commenting on the prospects for hybrid and regenerative technology. The barriers to automation, in particular labour unions, cost, inflexibility, and problems of software integration are also analysed. The presentation culminates with an assessment of the future for automation as the shipping industry emerges from the current economic crisis.

TRENDS IN OPERATIONAL TRANPORTATION PLANNING FOR ROAD CARRIAGE AND PROBLEMS TO BE SOLVED

Herbert Kopfer

Chair of Logistics
University of Bremen
Wilhelm-Herbst-Straße 5
28359 Bremen, Germany
E-Mail: kopfer@uni-bremen.de

1. INTRODUCTION

In this talk an overview of the typical problems to be solved for operational transportation planning will be presented and it will be shown how these problems and the corresponding models changed over time. At first the original Traveling Salesman Problem (TSP) and the first version of the Vehicle Routing Problem (VRP) are presented. Then a brief outline of classifications for vehicle routing problems is presented. This outline focuses on the problems which have been investigated most intensively in literature during the last decades. Subsequently, the characteristics of typical problems arising in the wider area of operational transportation planning are discussed and it is demonstrated how the planning tasks evolved from the simple original TSP and VRP to more complicated and comprehensive problems which are investigated in current research and applications.

2. COMPONENTS OF THE OPTIMIZATION MODELS

The objective functions of state-of-the-art models for operational transportation planning problems evolved from the primal goal of minimizing the sum of the driven distances of all vehicles to more realistic and complex goals. These are for instance the minimal number of used vehicles, minimal operation time, minimal operations costs, maximal flexibility, maximal reliability, minimal fuel consumption, minimal pollution, maximal environmental sustainability and last not least more complicated compound multi-objective goals referring to costs, service levels and ecology. The restrictions for the models representing vehicle routing problems changed from (a) a set of simple conditions for visiting all customers, route continuity and capacity limitations over (b) the introduction of time aspects, e.g. by windows and even multiple time windows to (c) complex descriptions of the solution space representing complicated situations with important restrictions which should not be ignored like e.g. drivers' driving time limitations and break scheduling.
In conventional approaches, the construction of routes and schedules has been restricted to vehicles performing the transportation tasks. The planning data for the goods to be transported have been considered as fixed and given in advance. In some progressive approaches the set of objects for which plans have to be derived (i.e. objects which are to be routed and scheduled) is extended to the cargo or to the used loading equipment. Loading equipment becomes more and more expensive and it often is a scarce resource which itself must be planned for transportation in order to have it available in spite of given imbalances. That is why resource planning for loading equipment becomes more and more important. In some approaches cargo can autonomously plan its way from its origin to its destination using a given infrastructure for transportation services. Of course, in these scenarios a matching of the cargo's planning and the planning for the vehicles providing the needed services is necessary. In other approaches, containers or swap bodies are objects of routing, too. With respect to routing, containers and swap bodies are passive objects since they cannot move without being carried by a vehicle which takes over the role of an active object. Thus, the routes of the used vehicles and the used containers must be matched in such a way that the routes of the different types of objects are optimized and each container movement is assigned to a specific vehicle transportation task.
Originally, the task of vehicle routing and scheduling concentrates on the optimization of the entirety of all constructed routes for a given set of transportation requests and a given (maybe capacity-limited) set of transportation resources. In this case, the planning process is oriented towards a given optimization criterion which usually is related to the degree and amount of used transportation resources. In addition to that original task, some advanced approaches try to arrange the routing and scheduling in such a way that different routed objects are synchronized; e.g. different vehicles are synchronized because they have to exchange swap bodies at a predefined location or in scenarios of container

scheduling there must be a synchronization between containers and vehicles because for each leg of its route each container needs a vehicle assigned to the resulting transportation task.
Finally, the scope considered for generating transportation plans is extended form the fleet of a single enterprise to the more challenging problem of harmonizing the planning of several enterprises being partners in a collaborative planning process. Such collaborative approaches allow the partial exchange of transportation requests among the partners and thus offer the chance for finding improved solutions with an additional collaboration profit. Examples for challenging research tasks in the area of collaborative transportation planning are the definition of suitable and reliable exchange mechanisms and the development of fair and sustainable profit sharing models. Some approaches even go beyond the before mentioned type of collaboration (exchange of transportation requests) by applying a concept for sharing transportation resources. In particular, models for sharing containers among several partners might be a remedy to reduce empty container transportation.

3. CURENT RESEARCH FIELDS

In general, there are several problem fields in the area of operational transportation planning which arise from the above trends for the evolvement of models for vehicle routing and scheduling. These fields refer to the following topics:

- Green transportation,
- Integration of self-fulfillment and sub-contraction,
- Synchronization of routes; e.g. between several vehicles or between active and passive transportation objects,
- Combination of vehicle routing problems with related problems; e.g. combining vehicle with location problems (location-routing problems) or combining routing and flow problems (in hinterland transportation planning or the integration of self-fulfillment and sub-contracting),
- Collaboration in groupage systems for request exchange, exchange of loading space, (temporary) sharing of vehicles, container sharing.

The above problem fields constitute exciting and challenging research tasks. These problem fields will be presented and shortly discussed in this contribution. Additionally, some important chances and obstacles for the research in these fields will be characterized. Some selected fields will be treated more intensively by introducing and discussing illustrating examples.

ACKNOWLEDGEMENTS

The author is grateful to the generous support from the German Research Foundation and from the Korean Research Foundation.

Session B1: SCM 2

·**Day1: Sep. 15 (Wed.)**

·**Time: 13:20 - 14:40**

·**Chair: Marcel W. Ludema**

·**Room: Camellia, 5F**

LOGMS

The 1st International Conference on Logistics and Maritime Systems

September 15th–17th, 2010 Busan, Korea

Lin

Challenges of Healthcare Supply Chain Improvement in Thailand

Duangpun Kritchanchai [1] , Sineenart Krichanchai [2]

[1] Department of Industrial Engineering,
Mahidol University
25/25, Phuttamonthon 4 Road
Salaya, Phuttamonthon, Nakhon Pathom 73170 Thailand
e-mail: egdkc@mahidol.ac.th

[2] Centre of Logistics Management,
Mahidol University
25/25, Phuttamonthon 4 Road
Salaya, Phuttamonthon, Nakhon Pathom 73170 Thailand
e-mail: mai_sk158@hotmail.com

Abstract: Logistics and Supply Chain management has become a competitive tool in many industries. Supply chain perspective has led the industry to see through the process integration from upstream to downstream. In Thailand, healthcare supply chain has been regarded partially in hospital. The supply chain concept has just been introduced to this industry at an early stage. It is evidenced that healthcare industry in Thailand still suffers from inefficient process, inconsistent and inaccurate data information. Lacking in transparency throughout the chain makes it difficult to track and trace patient data and product data. Furthermore, each player in healthcare supply chain, manufacturers, distributors, healthcare providers, tends to develop its own data language. Therefore it discourages the player to share information with its partner. The lack of integrity in supply chain has become a facilitating factor for the growing problem of counterfeiting. This problem not only leads to economic loss but also has a negative impact on patient safety in Thailand.

The purpose of this paper is to review problems and challenges in healthcare arena regarding to supply chain management. It is found that standardized coding, operational re-engineering and Implementing information technology are promising performance improvement. The paper develops a framework for healthcare supply chain improvement in Thailand. It urges both policy makers and implementers to realize national health problems as well as identify rooms for further research and improvement.

1. INTRODUCTION

Logistics and Supply Chain management has become increasingly important in recent years as supply chain perspective has led the industry to see through the process integration from upstream to downstream. Due to this strategy, it enhances trading partners to share their information and cooperate to improve efficiency of the supply chain. Effective coordination along the chain has played an important role in focusing on the innovation, flexibility and speed that bring about the competitive advantage necessary for survival in a competitive business world (Turhan & Vayvay, 2009).

Supply chain management is more complex in healthcare industry. In this area, companies and hospitals have to do highly accurate job as cost of error might be someone's life. (Mustaffa & Potter, 2009; Turhan & Vayvay, 2009)In Thailand, healthcare industry has been regarded partially in hospital. Operations and co-ordinations across players have been neglected. The supply chain concept has just been introduced to this industry at an early stage. It is evidenced that healthcare industry in Thailand still suffers from inefficient process, inconsistent and inaccurate data information. This idea is supported by one of the biggest state-owned hospital's case study in Thailand, the case study shows that there is a lack of proper inventory management and an inefficient internal supply chain (Kritchanchai and Suwandeechochai, 2010).

1.1 Supply chain management in healthcare industry

Gattorna (1998) describes a healthcare business as it is provided by a variety of product and service enterprises including medical consumables, pharmaceuticals, catering, laundry cleaning, waste management, home-care products, information technology, vehicle fleet management and general supply. A typical supply chain is a complex network consisting of many different parties at various stage of the value chain (Kritchanchai & Suwandeechochai, 2010; Rossetti, 2008 Mustaffa & Potter, 2009; Turhan & Vayvay,2009; Burn, 2002). Generally, healthcare supply chain has similar core structure likewise other industries' supply chain as it composed of input, process and output. There are material flow and information flow along the chain (Kritchanchai & Suwandeechochai, 2010). Under the concept of supply chain management, the merchandises are produced and delivered to the right quantities, at the right location and at the right time (Rossetti, 2008).However, Turhan and Vayvay (2009) argue that it is a must in healthcare industry as a cost of error might be someone's life

According to Mustaffa and Potter (2009) and Burns (2002), the four major types of players are Manufacturers, Distributors, Healthcare providers and Payers. The typical supply chain can be found in Figure 1. Manufacturers include primary and secondary manufacturers. In term of primary manufacturer, this refers to any manufacture that involves the creation of active ingredient contained within the medications. Primary manufacturers act as a supplier for second manufacturer. After obtaining active ingredients from primary manufacturers, second manufacturers are responsible for transforming active ingredients into a useable medicine such as capsules, tablets, solution and so on. It can be said that manufacturer section influences on the pharmaceutical prices, assessing expected demand, future competition and project marketing costs.

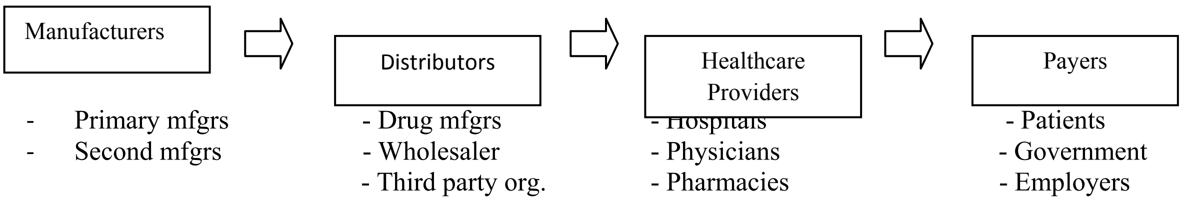


Figure 1. Healthcare supply chain (Adapted from Mustaffa and Potter (2009) and Burns (2002))

The finished products then are distributed to healthcare providers by distributors, wholesalers or manufacturers itself. Instead of allowing a third-party to deliver their products, there are many manufactures that are responsible for distributing their own products allowing them to leverage margin on self-manufactured products to discount the distribution fee. This might solve the problem caused by the third party distributors deliver the substitute products to the hospitals. As it is found that in several cases the third-party distributors cut its inventory stock of the competing product due to reducing cost of stock control.

2. LITERATURE REVIEW

2.1 Current problem

Healthcare industry has suffered from poor quality information which contributes ineffective healthcare services and also negatively affects the outcome of treatment for patients (Gibbons, 2009).This is supported by the studied in Shanghai. It found that three major aspects have caused traceability problems in Implanted Medical Devices (IMD). Firstly, information recorded during the process was inaccurate. Secondly, Manufacturers hardly collect the actual use data from hospital so that they are not able to fulfill their post-market responsibilities. Thirdly, the information of IMD use was not transparent so that it is really hard to protect patient safety, rights and interests (Yan, 2009). Inconsistent and inaccurate product information negatively impacts the rest of healthcare supply chain in the U.S as well. Each year more than $11billon of waste has been spent due to inefficient process, order and invoice error and outdated information technology (Pleasant, 2009).The difficulty faced by NHS, healthcare provider in UK, is that the lack of data standards contributes to many data silos. Data is available in a large volume whereas quality information is in short supply. From the NHS database, it contains 130 different descriptions of one single product. Due to this problem, it means that the analysis of expenditure and demand requirement in term of time and resources across organizations is very costly in term of time and resource (Gibbons, 2009). Gibbon(2009) highlights that without standards to identify the products and supplier, the accuracy cannot be certain and the transparency is limited.

Worse, healthcare supply chains exist as highly fragment systems in which manufacturers, distributors, wholesalers and providers operate independently from one another. Fragmentation complicates the task of connecting the thousand of partner involved at any stage in the chain (Dobrzykowski & Vonderembse, 2009; Burns, 2002). Burns (2002) also addresses that all parties in healthcare industry still lacks coordinated effort and knowledge sharing. Gibbons (2009)states that healthcare is an information intensive environment and the availability of quality information is essential for the delivery of safe and effective healthcare services . Due to lack of information sharing the supply chain acts more to push products down the chain rather than pull them from the customer. As a consequence, each stage holds inventory to prevent stockouts and providers order products based on just-in-case inventory planning (Burns, 2002).

Moreover,it i s apparent t hat inefficient process leads t o poor perform ance and avoi dable cost i n heal thcare industry(Kumar, 2008).The Glaxosm ithkline, a pharm aceutical company, found that the traditional custom er managed lead to the inability to meet changing demand pattern and increased transportation costs due to inefficient planning and difficulties for the supplier to determining production capac ity (Danese, 2004). Guy's & St. Thomas hospital faced difficulty because the traditional system of hospital pharmacy has inherent effici encies and recruitment and retention problems wi th the pharmacy professi on. Due to this problem, it affects directly to hospital servi ce and pat ient safety (McRobbie et al.,2003) The complicated role of professional healthcare also impacts on hospital process. Owing to the diverse management structure in hospital, providing co-ordinate care was difficult (Nicholson, 1995). It can be seen that three groups of probl ems are apparent . There are i nefficient busi ness process, dat a i nconsistency and fragm ented system.The table 1 illustrates the problem found in healthcare industry.

Table 1. Problems in healthcare industry

Studies	Highlighted key finding	Problem categories
Nicholson (1995)	Owing to the diverse management structure in hospital, providing co-ordinated care was difficult	Inefficient business process
McGrath and More (2001)	Poorly integrated Information system	Data Inconsistency and Fragmented system
Burns (2002)	All parties in healthcare industry still lacks coordinated effort, strategic alliance formation and knowledge sharing	Data Inconsistency and Fragmented system
McRobbie et al. (2003)	Inherent efficiencies and recruitment and retention problems within the pharmacy profession impacts on healthcare service and patient safety	Inefficient business process
Danese (2004)	The traditional customer managed inventory leads to the inability to meet changing demand pattern and increased transportation costs	Inefficient business process
Kumar et al. (2008)	Increasing operating cost is caused by the inefficient centralized supply system and material management	Inefficient business process
Dobrzykowski and Vonderembse (2009)	A fragmented system in healthcare industry limited communication among all parties	Fragmented system
Gibbons (2009)	Poor quality information leads to ineffective healthcare services and negatively affects the outcome of treatment for patients	Data inconsistency
Pleasant (2009)	In the US, healthcare industry suffers from inconsistent and inaccurate product information	Data inconsistency
Yan (2009)	Three major aspects have caused traceability problems in Implanted Medical Devices (IMD)in Shanghai	Data inconsistency

2.2 Enhancing performance improvement and supply chain integration

Regarding to the current situation, we conducted a literature review in order to find out how health care enterprises adopt the strategy to alleviate these problems and improve their operational performance. To promote healthcare supply chain improvement, st andardization has been m ajor concern. Im plementation of st andardization cont ributes t o information sy nchronization so t hat all st akeholders are able to speak t he sam e electronic l anguage (Krey sa and Denecker, 2009). According to European Federation of Ph armaceutical Industries and Associations's report (2008), European countries have experienced in im plementing standardization. At the initial stage, it mainly used by manufacturers, distributors and pharmacists for l ogistical propose whereas, t oday, it i s al so widely used for administrative reasons as well (EFPIA, 2008). In 2007, A premier healthcare industry provider, an American healthcare industry leader, accelerate the use of a single unique device identif ication (UDI) system in healthcare industry by proposing the requirement to their supplier to incorporate certain standard in order to win contracts(Pleasant, 2009).

Moreover, Shanghai Food and Drug Administration began to implement the IMD traceability system in November 2006 due to post-market surveillance purposes. The tracking system covers more than 100 hospitals using IMDs in Shanghai and the IMDs included high-risks devices such as orthopedic internal fixation devices, orthopedic implants, synthetic crystals, breast implants, and pacemakers(Yan, 2009). NHS trusts in UK and its suppliers have experienced the implementation of common data standard across the procurement and commercial systems. This enables interoperability between systems, allowing automation which reduces the resources required, removes errors, increases compliance and reduces risk (Gibbons, 2009).

Gibbons (2009) highlights that more effective process and better relationships provide a higher quality supply chain. To enable the effective delivery of high quality information organizations on the demand and supply sides of the healthcare, network must be able to share data. Currently, Healthcare providers attempts to integrate upstream with the manufacturers, wholesalers and distributors (Rossetti, 2008). After implementing procurement eEnablement technologies provide the NHS with an important opportunity to enhance its capacity to manage procurement information, improve its commercial and procurement processes and remove waste and duplication (Gibbons, 2009

Besides this, healthcare providers understand information and communication has a potential for improving delivery and quality of care. Therefore they implement information technology so as to share information to their patients. (Dean, 2003). In Italy, the development of the new national healthcare information systems (NSIS) has taken place in the context of gradual change. The NSIS would become a "connectivity backbone" between regional IT systems which empower the SSN (Servizio Sanitario Nazionale-National Healthcare service) more efficient and also delivering better services to the individual (Bergamachi, 2003). Kaiser Permanenete-the largest non-profit health maintenance organization in the USA with 8.4 million members nationwide which signed a 1.6 billion euro contract in early 2003 for a highly pervasive EHR program (Mori, 2003).

In healthcare industry, business process redesign is used to implement organizational transformations towards more customer-focused and cost-effective care. McRobbie et al.(2003)studied the introduction of pharmacy service near patient ward to simplify discharge processes. Re-engineering pharmacy service resulted in efficiency of service, more appropriate use of skill mix and the quality use of medicines. Derbyshire Royal Infirmary, a trust in UK, launched a business process re-engineering project. Accident and emergency was selected as a pilot site. This project focuses on Changing the role –specific culture to create more flexibility and to be more responsive to sudden change in workload (Nicholson, 1995).

Table 2. Inventions of healthcare supply chain improvement

Study Highlight	key finding	Intervention
Nicholson (1995)	Focused on Changing the role –specific culture to create more flexibility	Business process re-engineering
Bergamachi (2003)	In Italy, the development of the new national healthcare information systems (NSIS) has taken place in the context of gradual changeAcademic Pediatric Practice.	Information sharing
Study Highlight	key finding	Intervention
McRobbie et al.(2003)	Studied the introduction of pharmacy service near patient ward to simplify discharge processes.	Business process re-engineering
Mori (2003)	Kaiser Permanenete-the largest non-profit health maintenance organization in the USA implemented EHR program	Information sharing
European Federation of Pharmaceutical Industries and Associations's report (2008)	European countries have experienced in implementing standardization.	Standardization and information sharing
Rossetti (2008)	Healthcare providers develop strengthen relationship and integrate vertically upstream with its suppliers due to an increasing in material and supply costs	Information sharing
Gibbons (2009)	The implementation of data standard across the procurement and commercial systems by NHS trusts in UK and its suppliers enables information to easily transfer between systems.	Standardization and Information sharing

Kreysa and Denecker (2009)	The Implementation of GS1 standards in healthcare industry	Standardization
Pleasant (2009)	The Premier healthcare alliance implement global standard in order to improve its operational performance	Standardization and information sharing

3. FOCUSED GROUP

To enhance healthcare supply chain improvement in Thailand, several workshops have been conducted, under the project of National Research University. In the workshops, we invited various players in healthcare industry including hospital's board of directors, pharmacists, system engineers, private supplier and distributors and expertise from other industries. The purposes of these workshops were to find out the way to enhance performance improvement and supply chain integration in healthcare. National standard has been one of concerned issues among trading partners in the industry. Health care industry involves a complicated network; including manufacturers, distributors, healthcare providers and reimbursement institutions. Each player tends to develop its own code in order to succeed particular goal. While manufacturers and distributors develop the code for the management of trade flows and logistical purposed, healthcare providers adopt a codification system public health reason. Even though these enterprises are in the same chain, they speak different language which affects to the flow of information along the chain. Worse, it is impossible to select one of these codification systems for and implement to all players in the industries. Since the codification is created to particular player it still lacks in essential information for another player. This has raised the question whether healthcare industry in Thailand should develop the national standard for product identification or select the existing global standard to implement. Therefore, all players in the industry would speak the same language.

Additionally, it can be seen that healthcare industry is a fragmented industry as it consists of a large amount of parties. All this parties tend to operate independently without coordinated effort or any particular concern on alliance formation. Notwithstanding, to enhance integration healthcare supply chain, it requires trading partner to improve business relationship and share more information. Information sharing among trading partner taken into account However, there are two challenges that should be concerned in the healthcare industry before sharing information which are; confidentiality and system security. Regarding to confidentiality, it should be considered the extent of which information will be shared to the supplier in order to improved material management and which information that are not allowed to be shared due to the policy or the legislation. Another challenge is the system security. Both trading partners has to decide the level of integration and then considerable effort goes into identifying technologies that provide strong system defenses against external attack and given it is healthcare, the costs are justified.

Business process re-design is another issue that various organizations have taken into account. In the focus group, we have discussed why firms tend to implement this "Change" strategy instead of carry on the former strategy. Problems identified include customer dissatisfaction, difficulties with current material management and cost that higher than necessary. Therefore healthcare enterprises found that the former strategies were not able to deal with these problems. As a consequence, they decides to implement a business process re-engineering strategy and aim to achieve dramatic improvement in critical performance measurement such as cost, quality, service and speed. Additionally, the use of information technology is conceived as a contributing factor that enhances firms seeking for the change of structure organization. It can be concluded that the business process re-engineering is necessary in order to improve business efficiency and responsiveness.

4. PROPOSED FRAMEWORK

As mention above, we conducted the literature review so as to find out the current problems occurred in the healthcare industry and the adopted strategy for alleviating all these problems. Regarding to the literature review, we can conclude that there are three mains problems which are inconsistent and inaccurate information, fragmented system and inefficient business process. Due to these problems, they pose threat for all players in healthcare industry. Thus, healthcare enterprises have sought several strategies to deal with these problems. As a consequence, it comes up with three interventions; standardization, information sharing and business process re-engineering. Firstly, implementation of set of standard and system enable suppliers and buyers to identify and navigate the product along the chain. Therefore this intervention ensures correct products are delivered to correct locations, leading to an increase in patient safety. Secondly, healthcare enterprises tend to implement information technology in order to facilitate product information to their trading partner. Benefit gained from information technology is to enable traceability and contribute to improvements in patient safety. Thirdly, business process re-engineering is also a recommended intervention when the former strategy fails to meet the performance measurement; cost, quality, service and speed. Therefore, the change strategy is able to enhance process efficiency and responsiveness.

Regarding to the focused group, it also supports the idea from the literature review. It can be concluded that the need for accurate and consistent accelerates implementation of the data standard and system. Therefore healthcare enterprises have to put an effort to develop the uniform product identification. It might be to develop national product identification for healthcare industry in Thailand, otherwise, implement an existing global standard. Additionally, traceability is another concerned. The ability to track and trace product allows the healthcare enterprise to enhance service improvement and error detection effectiveness. However, to improve traceability, it means healthcare enterprises have to share information with their trading partner. According to this point, Establishing trust and strengthen relationship is necessary. Besides this, the process efficiency and responsiveness is the reason why healthcare enterprises implement the business process re-engineering strategy. The former strategies are no longer to manage business process effectively so that it accelerates the adoption of the change strategy.

As mentioned earlier, data gained from both literature review and focus group provides a clear picture of factor enabling healthcare improvement. This is also embedded in the proposed framework. An illustrated of the framework is shown in figure 2. In this framework, there are three main enabling factors to the achievement which are identification, traceability and process efficiency and responsiveness. In term of identification, any materials flows, including medicine, medical device or even patient, along the supply chain have to be identified. Besides, any material flows and data flows have to be tracked and traced at any stage in the supply chain. The last factor is process efficiency and responsiveness. This means any process at any stage need to run effectively and respond abruptly. As can be seen in the figure, we suggest a three-dimension comprehensive intervention strategy to bring about the healthcare supply chain improvement in Thailand; internal process improvement, supply chain identifications and traceability and electronic Health Record.

Regarding to internal process improvement, the target groups of this intervention are both public and private healthcare provider. This intervention aimed at improving any process in hospital including warehouse management, inventory management, product identification, medication error and so on. At this level, it requires simple information technology in which the hospital is able to develop by its own. In term of supply chain identification and traceability, the target groups of this intervention are all players along the supply chain. The intervention focuses on enhance integration from the demand to the supply side. This means any material flows, from the manufacture to the patient at the hospital, is able to be identified and traced back at any point in the chain. At this level, it requires all players to implement the same data standard and system so that they can speak the same electronic language. By considering Electronic Health Record, this focuses on Nationalwide supply chain. It requires government regulator to enforce and implement national data standard such as health record, Medical device ID and so on. Therefore, healthcare enterprises in Thailand are able to speak the same language which contributes to patient safety.

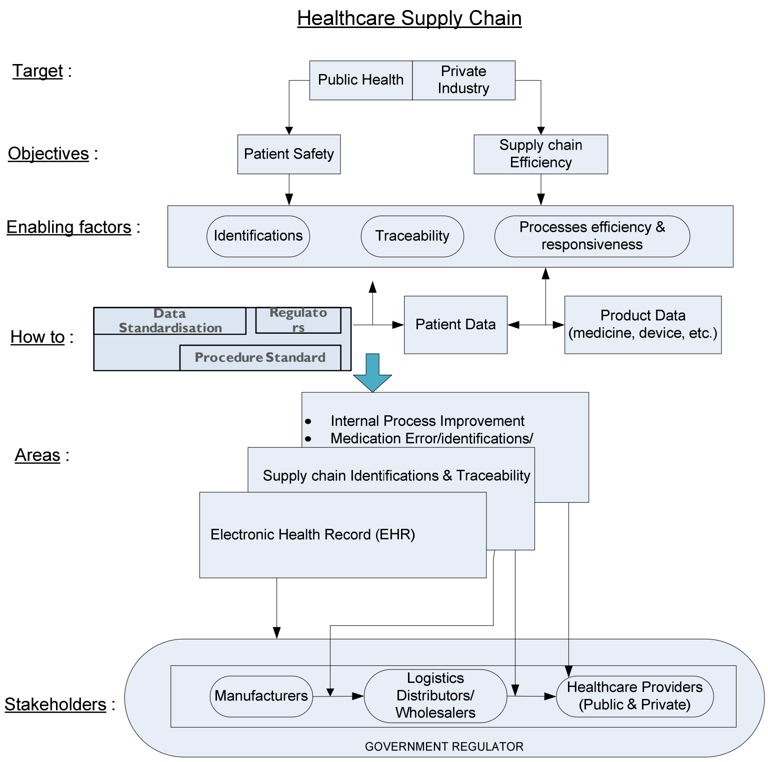


Figure 2. The framework for healthcare supply chain improvement in Thailand

5. DISCUSSION AND CONCLUSION

We developed the framework for healthcare supply chain improvement in Thailand by first exploring the current problems occurred in the industry and the intervention. We found that, in Thailand, healthcare supply chain has been regarded partially in hospital. Operations and co-ordinations across players have been neglected. It seems that most hospitals focus only on healthcare performance and undermine its own operational performance. However, supply chain problems, data inconsistency, fragmented system and inefficient business process, are not only affect operational performance but also leads to negative impacts to patient safety. Due to these problems, it is no longer to solve individually but it requires all players in the industry to start implementing supply chain thinking and collaborate with their partners to alleviate the problems. However, in Thailand, it is apparent that the supply chain concept has just been introduced to this industry at an early stage. It can be said that the proposed framework is the first step to urge all healthcare players and policy makers to aware of these problems.

There is still more room for research in developing countries. It is apparent that a number of research proposed various interventions, aiming at alleviating healthcare supply chain problems, are concerned the problem occurred in the developed countries which might not suit for the situation in Thailand. Therefore, instead of following this intervention, We tried to find out the why healthcare enterprises exactly need from these interventions. We found that three main concepts behind these interventions are identification, traceability and process efficiency and responsiveness so that we embedded these concepts as enabling factors for healthcare supply chain improvement. Thereafter, we developed framework and intervention based on these concepts and situations in Thailand.

6. CITING REFERENCE

Burns, L.R., DeGraaff, R.A., Danzon,P.M., Kimberly,J.R., Kissick, W.L. and Pauly, M.V. (2002). The Wharton school study of health care value vhain. In L.R. Burns (Eds.)*The Health Care Value Chain.*(pp.3-23). Retrieved from media.wiley.com/product_data/excerpt/17/.../0787960217.pdf –

Bergamashi, W. (2003). Basic concept model of the new national healthcare information system (NSIS). In K.Dean(Eds.), *Thought leaders: Essay from health Innovators* (pp.72-79). Retrieved from: www.cisco.mn/web/about/ac79/docs/wp/ctd/CISCO_Health_0317.pdf

Danese, P. (2004). Beyond Vendor Managed Inventory: the Glazosmithkline case. *Journal of Supply Chain forum,* 5(2), 32-40

Dean, K. (2003). Introduction-Essays from health Innovators. In K.Dean(Eds.), *Thought leaders: Essays from health Innovators* (pp.8-17). Retrieved from: www.cisco.mn/web/about/ac79/docs/wp/ctd/CISCO_Health_0317.pdf

Department of Healthcare (2000). *An organization with a memory: Report of an expert group on learning from adverse events*. Retrieved from http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4065086.pdf

Dobrzykowski, D.D. & Vonderembse, M.A. (2009). *Healthcare supply chain and IS strategy for improved outcomes.* Paper presented at the POMS 20th Annual Conference.

EFPIA (2008). *Towards safer medicines supply: A vision for the coding and identification of pharmaceutical products in Europe*. Retrieved from ec.europa.eu/enterprise/.../**pharmaceutical**s/.../114_efpia_en.pdf

Gattorna, J.L. (1998). *Strategic Supply Chain Alignment: Best practice in Supply Chain management*, Gower,Aldershot.

Gibbons, R.H. (2009). The NHS Procurement eEnablement program using information to deliver better healthcare. *GS1 Healthcare Reference Book 2009/2010*, 22-25

Kreysa, U. and Denecker, J. (2009). GS1 standards in healthcare: raising the bar on patient safety and supply chain efficiency. *GS1 Healthcare Reference Book 2009/2010*, 1-5

Kumar, A., Ozdamar, L. and Zhang, C.N. (2008). Supply chain redesign in the healthcare industry of Singapore. *Journal of supply chain management,* 13(2), 95-103

McRobbie, D., Badnall, R. and West,T. (2003). Assessing the impact of re-engineering of pharmacy services to general medical wards. *Pharmaceutical Journal,* 270(7239), 342-345

Mori, A.R. (2003). Cooperative development of the healthcare infostructure for Europe. In K.Dean(Eds.), *Thought leaders: Essays from health Innovators* (pp.90-103). Retrieved from: www.cisco.mn/web/about/ac79/docs/wp/ctd/CISCO_Health_0317.pdf

Mustaffa, N.H. and Potter, A. (2009). Healthcare supply chain management in Malaysia: a case study. *Journal of Supply Chain Management*, 14(3), 234-243

Nicolson, J. (1995).Patient-focused care and its role in hospital process re-engineering. *Journal of*

Health Care Quality Assuarance, 8(7), 23-26

Pleasant, J. (2009). Changes has finally come: U.S. Healthcare industry to implement common data standards to improve safety, reduce cost. *GS1 Healthcare Reference Book 2009/2010:* 6-9.

Rossetti, M. D. (2008). *Inventory Manage Issue in Healthcare Supply Chains.* Retrieved from www.uark.edu/~rossetti/reports/**healthcare_supply_chain**_rep.pdf

Salam, M.A. (n.d.) *Value creation through lean healthcare*(Unpublished Thesis). Assumption University, Bangkok, Thailand.

Shah, N. (2004). Pharmaceutical supply chains: key issues and strategies for optimization. *Computers and Chemical Engineering,* 28, 929-941

The Health Startegies Consultancy LLC (2005). Follow the Pill: Understanding the U.S. Commercial Pharmaceutical Supply Chain. Retrieved from www.kff.org/.../Follow-The-Pill-Understanding-the-U-S-Commercial-Pharmaceutical-Supply-Chain-Report.pdf -

Turhan, S.N. and Vayvay (2009,July). *Modeling of VMI implementation via SOA in a Healthcare Supply Chain.* Paper presented at the European and Mediterranean Conference on Information Systems. Retrieved from www.iseing.org/emcis/.../Proceedings/Presenting%20Papers/C66/C66.pd

SUPPLY CHAIN ENGINEERING; A SYSTEM OF SYSTEMS PERSPECTIVE

Marcel W. Ludema

Faculty or Technology, Policy & Management
Delft University of Technology
Jaffalaan 5
Delft, The Netherlands
e-mail: m.w.ludema@tudelft.nl

Abstract: To design supply chains means to incorporate the right requirements in the design. Without stating the right requirements it is impossible to know how stated requirements will lead to effective and efficient supply chain designs. This research identifies the supply chain as a system and as a system-of-systems. The system-of-systems approach is a relative new research discipline for defining, abstracting, modeling, analyzing and synthesis of systems and system-of-systems for both research and design problems. Supply chains are discussed and analyzed on several aggregation levels. Next a framework is developed to contribute to the effective and efficient design of new supply chain. For the assessment of this framework and to learn about necessary improvements the application of it during the design and engineering of supply chain in the professional world is advised.

1. INTRODUCTION

Systems thinking and the subject of 'system' aspire to the status of a serious academic discipline, (Checkland, 1993). Systems engineering is combining the disciplines necessary to create and build large construction projects many of them we use and can see around us. All of these projects were created by humans and societies. Projects are added in time or we can say that the societies they are part have grown step by step. So it can be said that humans have created societies by adding in due time and for several different reasons societal parts (projects) to these societal system. Often this creation process was not planned upfront but more the result of a process that evolved step by step.

Institutions are subsystems of these societal systems also added by humans to perform certain functions that fulfill specific demands. Some of these institutions are clusters of industries that in their decomposition are businesses that contain production-lines that facilitate the production of products that fulfill the needs a wide range of customers. A supply chain connects parts of several production lines of several firms (business to business) till finally the customer can be served (business to consumer). Supply chains can be seen as connectors between elements of several larger nested systems. These systems are complex containing other parts that are relevant for the functioning of these systems. In this paper the focus is on the elements that are relevant in understanding supply chains.

This paper is organized in the following way. First some general supply chain characteristics are discussed. Next systems engineering and system of systems engineering are explained. After this Systems Engineering and System of Systems Engineering are confronted with the concept of the supply chain. Finally conclusions are drawn and a recommendation for further research is given.

2. SUPPLY CHAIN CONCEPTS

2.1 Supply chains as demand fulfillment structures

A supply chain is: "A network of collective business entities responsible for procurement, manufacturing and distribution activities associated with one or more series of related products", (Coyle *et all*., 2003). The SCOR-reference model (Supply Chain Council, 2008), (figure 1) is a helpful tool to visualize and analyze supply chain processes that extend beyond the company level all the way to the suppliers' supplier and the customers' customer. The SCOR-reference model indicates organizational segments of the supply chain, not necessary belonging to the same company but responsible for the management and control of it. It is a useful tool to see if and how the coordination between these underlying processes is organized and what and where information sharing and communication takes place.

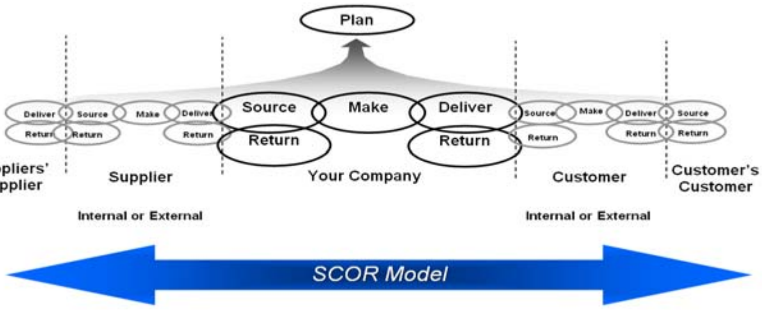


Figure 1: SCOR-methodology, (Supply Chain Council, 2008)

The infrastructure of supply chains is a collection of nodes and linkages to facilitate physical flows in such a way that an effective and efficient fulfillment structure is provided. Physical nodes are: points from where raw materials are sourced, factories and transfer centers (sea and inland ports, airports, warehouses, distribution centers and retail outlets). Physical linkages reflect the modalities that can be used to transport physical items between nodes. Besides physical nodes in the physical flow; supply chains contain informational, money (financial) and organizational nodes in the respective flows that all interrelate with each other.

In every supply chain a focal company controls the customer decoupling point; this is the point where the demand is dictated by the customer against a supplier organizing a fulfillment structure, (Hoekstra and Romme, 1992), figure 2.

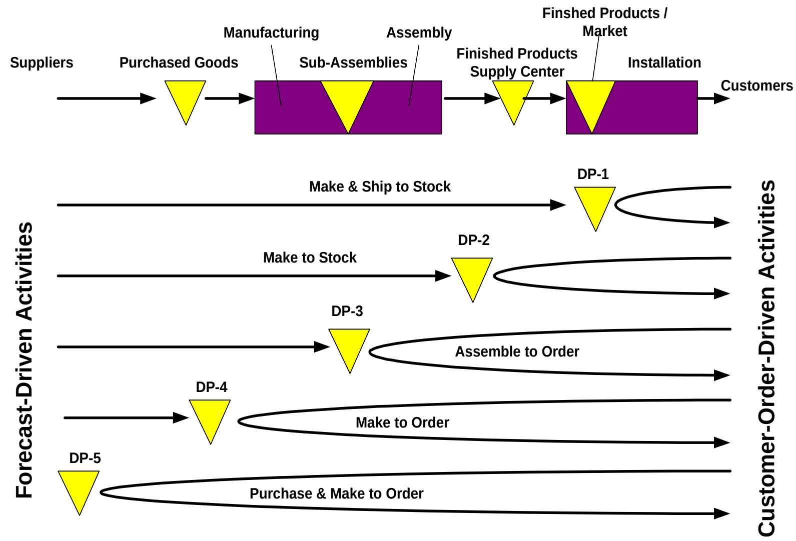


Figure 2 Customer order decoupling point; adapted from (Hoekstra and Romme, 1992)

2.2 Levels in supply chains and supply networks

Porter's value chain is a good way of communicating value-adding processes, (Porter, 2001). Transformation is transforming raw materials into assembly parts and assembly parts into final products. Stability is keeping and storing goods in facilities. Translation means goods in transition during transport between two locations. Possession is the exchange of ownership of goods and services. Sometimes goods are at the location of the potential buyer but not yet in the possession of this buyer. Basically in all four flows: physical, information and money value adding activities can take place, (table 2).

Table 1 Confrontations of value adding utilities

	Utility of Transformation	Utility of Stabilization	Utility of Translation	Utility of Possession
Organizational Flow	**Forward & Backward Integration / Segregation**	Keeping a Status Quo	Change Management	Buying / Selling of Organizations (take-over, buy-out)
Physical Flow	Manufacturing	Keeping Inventory	Transporting	Buying / Selling of Physical Products
Information Flow	Knowledge Creation (third order data / information use)	Data -Storage	Data Transmitting & Collection	Data Buying / Selling (Sharing) of Information
Money Flow	Creating Structured Financial Products	Keeping Cash Position	Paying & Getting Paid (Wiring Money)	Keeping Balance

The value network consists out of sequential and parallel organized value chains, (Porter, 2001). Several organizational relationships can be identified within these networks, either being currently existing organizational relationships or possible relationships. The standard is that all organizations are completely responsible of there own processes and only pure buying and selling relationships exists not sharing capacities and organizational responsibilities. The lines in figure 3 give some more possible relationships. Very common relationships are sharing purchasing processes, outsourcing value adding activities to suppliers and sharing the management of several layers in the supply chain.

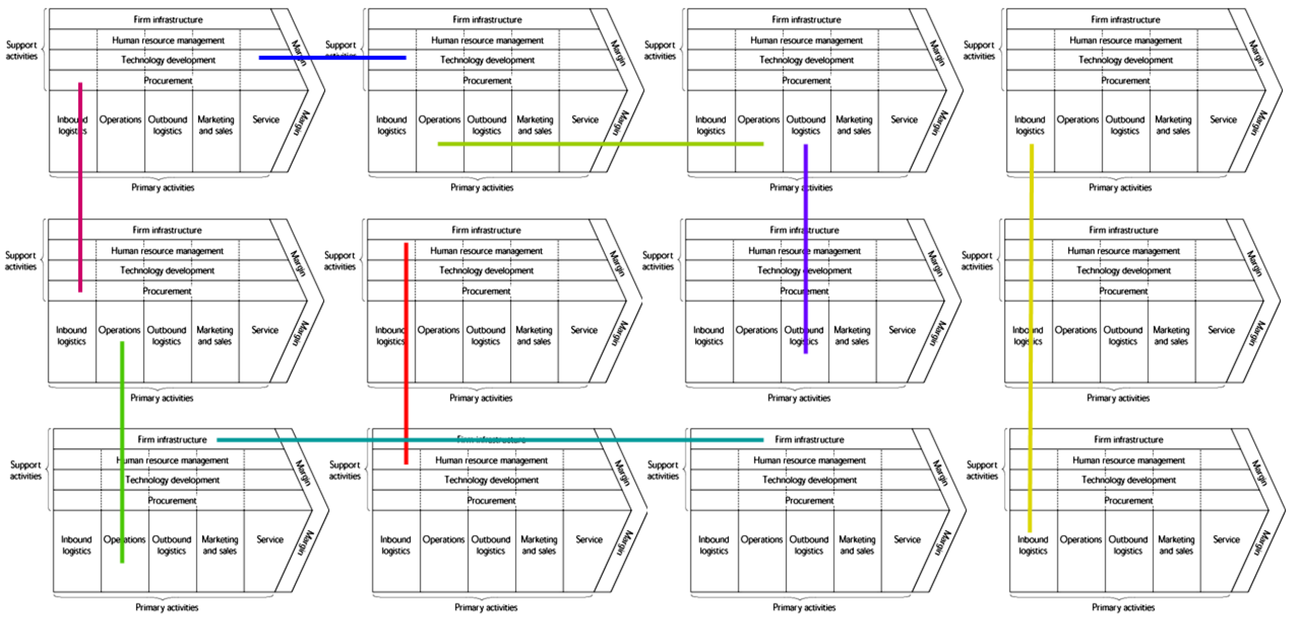


Figure 3 Organizational relations in a supply network; adapted from Porter (2001)

A supply chain can be decomposed in several levels, see figure 4, (Peck, *et all.*, 2002):

- Infrastructure and Asset Level (hard-ware);
- Process and Value Stream (flow-ware);
- Inter-organizational Network Level (org-ware).

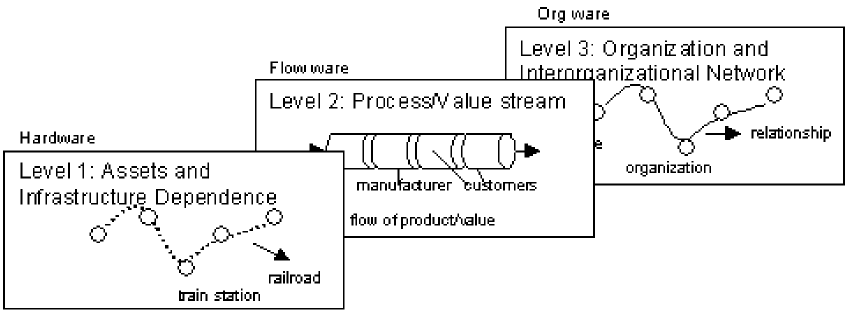


Figure 4: Levels of Supply Chain, (Peck, et all., 2002).

Value can be added in several ways. Some companies outsource non-core business activities. Often these activities are part of the physical layer of the value pyramid (figure 5). While outsourcing they concentrate on those activities that generate the most value. They find innovation in how to bring services to the market (virtual layer) and find that is more important that selling products alone. New combinations are made and knowledge of how to bring service to the market becomes crucial. Finally these companies realize value by using knowledge or using services of using products as their business. Within the design of supply chains these relationships need to be understood to understand where the real value is located and how several layers are interrelated.

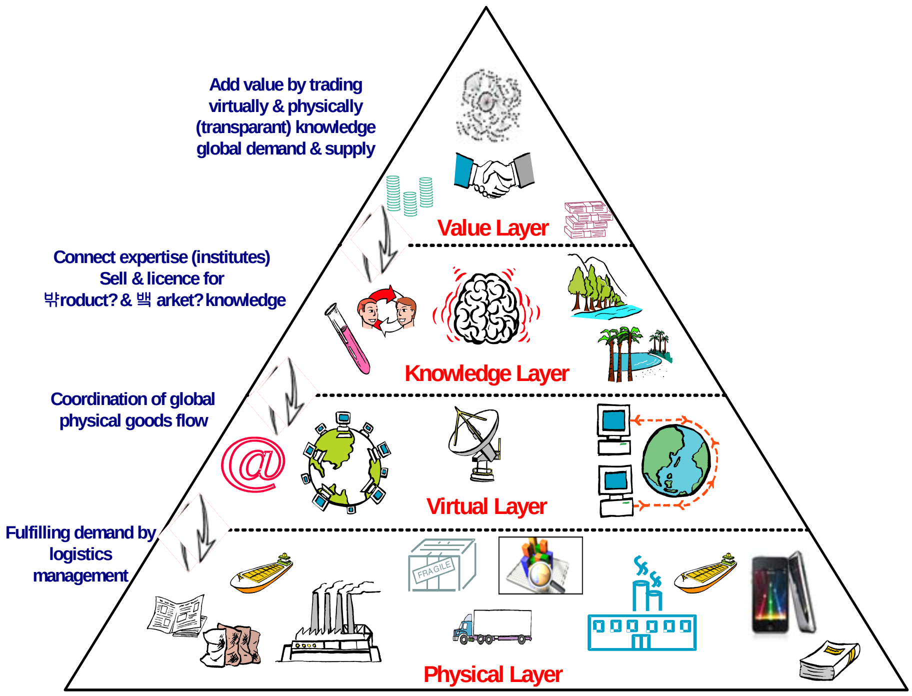


Figure 5: Multi-layer perspective on supply networks

3. SYSTEM APPROACHES, SYSTEMS ENGINEERING AND SYSTEM OF SYSTEM ENGINEERING

3.1 System approaches

Society can be visualized as a system that consists out of many subsystems e.g. cities but also industries and or individual companies that make it their business to facilitate bringing a wide set of services and products to a often wider set of markets. Humans can be considered as parts of organizations that in themselves can be considered parts of socioeconomic systems. Organizations, businesses, companies and institutions can be viewed as systems with the purpose to facilitate bringing services and products to other parts within these societies and finally to the people of these societies.

A system can be defined as: "a construct or collection of different elements that together produce results not obtainable by the elements alone", (Checkland, 1993). This is a very generic definition and looking at several application are as several different elements can be identified, like: people (operators and maintenance personnel), information (sub)-systems (hardware, software but also documents (manuals) and or schemes including operational instructions), facilities (for the main purpose of the system or for the support of it), strategies and policies (what guides the systems to produce its intended results), and many more things that contribute to system-level results. All man-made-systems are design or more formally and structured engineered.

3.2 System Engineering

In 1962 A.D. Hall developed the concept of a development to design and develop technical systems, basically the process of systems engineering (SE), (Brill, 1999), (Hall, 1962). Several elements where stated as important in the concept of SE. The first element is that, SE is multifaceted and can only be understood from different perspectives or facets. The second element relates to three important viewpoints that system engineers should use tackling problems that require SE. These three viewpoints are: (1) the physical or technical viewpoint, (2) the business or economic viewpoint, and (3) the social viewpoint. The third element considers the needs of its customers and determines how the needs can be met.

To understand the problem and the possible physical solution that solves the problem, the following questions should be answered: "(1) What is the system to do (performance, cost, time and reliability)? (2) What environment does it have to operate in (home, commercial, or military)? (3) What environment is the product to be made in (engineering/manufacturing skills and facilities)?" Blanchard and Fabrycky (1981) state that systems should be designed from a life-cycle perspective taking the full operational life of the physical solution into consideration. They state that the life-cycle of a system as "starting with the initial identification of a need and encompassing the phases (or functions) of: planning; research; design; production or construction; evaluation; consumer use; field support; and ultimate product phase-out".

SE is a very broad concept that related to the design and engineering in a team-effort of durable and complex system that have to be operated in a broad range of uses and situations. The field of SE is domain independent, though in several domains specific flavors of SE exist.

3.3 Systems engineering process decomposition

The System Engineering Process follows a set of phases going from establishing the need for a system, via stakeholder, consumer and user and technical requirement via a logical (functional) design and a structural (physical) design. The systems engineering process asks that each design has to be verified based on the stated requirements and validated after implementation, operation and maintenance. A wide set of sometimes more formal SE-approaches exists; in essence they contain the following phases:

- The first phase of the system engineering approach is focussed on the elaboration of a problem description, an opportunity statement and a mission statement. During this phase the problem at hand is analyzed in detail. Understanding the agreed upon problem is important to understand the contribution of the design and development process towards solving the problem at hand or that not solving the problem by adding something new is a better solution. This phase ends in a one or two sentence description giving the actual need of the solution to be developed; namely the need statement.
- The second phase is the analysis and development of the advanced planning process that guides and governs the system engineering process by means of mile-stone, project participants' involvement and stakeholder engagement proposals and deliverables.
- The third phase is the analysis of the expected operational profile of the to-be designed system. The fourth phase of the system engineering process is establishing a good set of stakeholder, consumer and user requirements. This is

established by interviewing stakeholders and future owners, users and maintainers of the system, by asking what they require from the newly to be designed system.
- The fifth phase is translating the stakeholder, consumer and user requirements into more technical system design requirements divided in functional requirements, non-functional requirements and program requirements. The quality and maturity of these requirements should be such that the system can be designed at the right system level.
- The sixth phase is the design of a concept description showing a first idea of the possible sub-systems and how the relate to each other and next to that the analysis and design of the context diagram indicating the interfaces of the to-be designed system with the environment and/or already existing systems.
- The seventh phase is the design of the temporal aspects of the system by using a state and mode analysis/design. Inputs to this phase are both the functional requirements and the operational profiles of the system.
- The eight phase is the design of the functional blocks of the system based on the functional requirements.
- The ninth phase is the design of the physical solution of the system and the allocation of the functional and temporal design with the physical design.
- The tenth phase is developing towards the integration of the temporal, functional and physical design and aligns them with the functional requirements. The concept of configuration management is also addressed in this phase, though it actually is a concept that needs attention earlier in the process from the development of the advanced system planning forth.
- The eleventh phase is developing the specification of one or more subsystems (because of time restriction in the course), towards a maturity so it could actually be constructed or manufactured.
- The twelfth phase is the verification of the design, basically by checking if how the fulfilment of the stated requirements is taken care of and how well they are incorporated in the final design.
- The thirteenth phase is the validation of the design; this is done by questioning potential users and owners of the system.
- The fourteenth phase is a reflection on the process followed.

3.4 System-of-systems approaches

Whether there is difference between what is considered a system and what a system -of-systems in under debate. The systems within systems can be seen as the Russian Babushka Dolls viewpoint. “You can start at the Universe and go all the way down towards quarks (and maybe beyond that) or from the quarks level all the way up to the Universe level.” This systems approach is useful to our thinking about man-made systems like international or national socioeconomic systems, industry clusters, separate industries, businesses, companies and production lines within these companies.

Taking a composition versus decomposition perspective and the different aggregation levels one can see the use of Systems Engineering at each of these aggregation levels or layers. On the following layers systems engineering can be applied, (Hitchins, 2007):
- Layer 5: Socio-Economic, where regulation and governmental control takes place.
- Layer 4: Industrial Systems Engineering, or engineering of complete supply chain/circles. Many industries together form a socio-economic system.
- Layer 3: Business Systems Engineering – many businesses together form an industry. At this level, systems engineering focuses on the optimization of performance often independent of other businesses within the industry.
- Layer 2: Project or System Level. Many projects together form a Business. Design and development and utilization of production lines or designing, developing and constructing complex artifacts falls under this layer.
- Layer 1: Product Level. Many products together form a system. This is the tangible artifact level. For many disciplines and their engineers this is considered “real” systems engineering.

Hitchins (2007) emphasize that this 5 layers model is a "nested" model. A set of products make a project, a set of projects make a business, a set of businesses make an industry and a set of industries make a socio-economic system. Kasser and Massie (2001) clearly state that the assumption of the 5-layer model are ”only approximate since a socioeconomic system has more in it than just industries, a business has more in it than just projects, actual organizations may divide the work in different ways resulting in either sub-layers, or different logical break points.”.

System-of-systems (SoS) are dedicated goal oriented systems that is a compilation of other dedicated goal oriented systems on a lower aggregation level. The SoS shares capabilities and resources on its lower aggregation level to achieve 'meta-system' capabilities and resources on a higher aggregation level. The SoS offers more functionality and performance than the sum of its component systems. “Systems of systems is a serious and evolving research discipline with a set of incomplete frames of reference, thought processes, quantitative analysis, tools, and design methods are incomplete.”, (Popper *et al.,* 2004). A number of SoS’s can be identified in this study. First durable systems that exist or are under design, second the engineering process that facilitates the creation of these system, third the decision making process that takes place during the engineering process. Fourth the knowledge creation and facilitation process that takes place during the decision making process that guides the engineering process of the systems.

The following five characteristics to distinguishing very large and complex but monolithic systems from systems-of-systems can be identified, (Maier, 1998): "

1. Operational Independence of the Elements: If the system-of-systems is disassembled in to its component systems the component systems must be able to usefully operate independently. The system-of-systems is composed of systems which are independent and useful in their own right.
2. Managerial Independence of the Elements: The component systems not only can operate independently, they do operate independently. The component systems are separately acquired and integrated but maintain a continuing operational existence independent of the system-of-systems.
3. Evolutionary Development: The system-of-systems does not appear fully formed. Its development and existence is evolutionary with functions and purposes added, removed, and modified with experience.
4. Emergent Behavior: The system performs functions and carries out purposes that do not reside in any component system. These behaviors are emergent properties of the entire system-of-systems and cannot be localized to any component system. The principal purposes of the systems-of-systems are fulfilled by these behaviors.
5. Geographic Distribution: The geographic extent of the component systems is large. Large is a nebulous and relative concept as communication capabilities increase, but at a minimum it means that the components can readily exchange only information and not substantial quantities of mass or energy. Minimum it means that the components can readily exchange only information and not substantial quantities of mass or energy."

So systems engineering and system-of-systems engineering are powerful concepts that are getting more and more attention as a general research discipline the field of supply chain management can possibly benefit from. So let's have a closer look at how supply chains and its related concepts, relate to the system-of-systems concept and SoS-engineering.

4. CONFRONTATION BETWEEN SUPPLY CHAIN PERSPECTIVES AND SYSTEM OF SYSTEM ENGINEERING

To understand if supply chains are systems of systems the characteristics as discussed before (in: Maier, 1998) have to be confronted with some characteristics of the supply chain. First it can be seen that systems engineering can be applied to the domain of the supply chain because the supply chain is clearly operating at all of the level of the 5-layer model of systems engineering: Socio-Economic, Industrial Systems Engineering, Business Systems Engineering, Project/System Engineering, and Product Engineering, (Hitchins, 2007). And following the characteristics of Maier (1998):

1. Operational Independence of the Elements: The supply chain can be disassembled into component systems that can operate independently but to created value they should be connected to other parts (in other systems). Within a supply chain network the independent supply chains can operate independently.
2. Managerial Independence of the Elements: Components of the supply chain can operate independently. One can buy elements like transportation or warehousing and suppliers and customers that can function as before or can be integrated in the overall supply chain. But also making the supply chain operate on the virtual level adds value to the supply chain.
3. Evolutionary Development: The supply chain is not fully formed. Its development and existence is evolutionary with functions and purposes added, removed, and modified with experience. The supply chain is operating in a dynamic environment has to adapt continuously and many example exist that it does.
4. Emergent Behavior: The supply chain provides functions that are more than the functions of its parts. The sum of all value adding activities provides a higher value than that of its components. These behaviors are emergent properties of the entire supply chain and cannot be localized to any component of the supply chain. The principal purposes of the systems-of-systems are fulfilled by these behaviors.
5. Geographic Distribution: The geographic extent of the component systems is large, transportation carries the supply chain globally and communication between the many components and between the many supply chains are frequent. Often communication between a sales and a purchasing department of different companies can be more frequent that the communication within the company itself.

5. CONCLUSIONS

This paper showed that the supply chain management can learn a great deal from the new and growing research discipline of systems engineering and system-of-systems (engineering) and that further research and attention seems worthwhile. Systems Engineering seems to be applicable for systems like supply chains. Systems-of-systems proves a valuable analysis approach to understand supply chain and learning from the qualities of systems engineering and system-of-systems engineering can benefit the design of future supply chains. It is clear that far more research is necessary to show the value of this answer.

6. REFERENCES

Checkland, P. (1993), *Systems Thinking, Systems Practice*, John Wiley & Sons, Chichester, USA.

Coyle, J., Edward J. Bardi, C. John Langley Jr. (2003), *The Management of Business Logistics, A Supply Chain Perspective, 7th Edition*, South-Western/Thomson Learning , Mason, Ohio, USA.

Supply-Chain Council (2008), *Supply-Chain Operations Reference-model, SCOR Overview, version 9.0*, Supply-Chain Council, USA.

Hoekstra, J.B., and Romme, J.H.J.M. (1992), *Integral Logistic Structures, Developing Customer-oriented Goods Flow*, Industrial Press, New York, USA.

M.E. Porter (2001), *Strategy and the Internet*, Harvard Business Review, March pp.62-78., USA.

Peck H., Abley J., Christopher M. (2002), Haywood M.M., Saw R., Rutherford C., Strathern M., Creating *Resilient Supply Chain, A Practical Guide*, Cranfield University, United Kingdom.

J. H. Brill, System Engineering (1999) – A Retrospective View, Journal of Systems Engineering. John Wiley & Sons, USA.

A.D. Hall (1962), *A Methodology for Systems Engineering*, Van Nostand, Princeton, New Jersey, USA.

B.S. Blanchard & W.J. Fabrycky (1981), *Systems Engineering and Analysis*, Prentice-Hall Inc. Englewood Cliffs, New Jersey, USA.

Hitchins, D.K. (2007), *Systems Engineering; A 21st Century Systems Methodology* Wiley, Chichester, USA.

Kasser, J. and A. Massie. (2001), *A Framework for a Systems Engineering Body of Knowledge*, 11th International Symposium of the INCOSE, Melbourne, Australia.

Popper, S., Bankes, S., Callaway, R., and DeLaurentis, D., (2004), *System-of-Systems Symposium: Report on a Summer Conversation*, July 21-22, 2004, Potomac Institute for Policy Studies, Arlington, VA., USA.

Maier, M.W. (1998), *Architecting Principles for System of Systems*, Systems Engineering, Vol. 1, No. 4, 1998, pp. 267-284., Wiley, Chichester, USA.

OPTIMIZING CRITICAL SUPPLIERS IN THE DEPLOYMENT OF A LEAN SUPPLY CHAIN

Mithun J Sharma[1, 2] and Yu Song Jin[2*]

[1] Centre for Management Studies
Dibrugarh University
Dibrugarh, India

[2] Department of Shipping Management
Korea Maritime University
Busan, South Korea

Abstract In a supply chain, the upper echelons plays a key role in tuning the chain performance. In the procurement cycle of the supply chain, the manufacturer outsources more parts and services to focus on their own core competencies. A situation may arise in which some suppliers may under perform in providing critical strategic supplies. This paper analyzes supplier performance using multi-criteria decision procedure along with Pareto analysis in identifying critical suppliers for supplier development to optimize their performance.

1. INTRODUCTION

In 2010, Toyota halted sales and production of 8 of their models that was included in a recall that concerns unintended acceleration caused by pedals. More than 2 million units were affected by this recall. In 2009, a recall that amounted to more than 4 million units was also made for the same issue but is said to have been caused by floor mat malfunction. The latest unintended acceleration issue, allegedly due to design flaw in the pedal, seems to have been made by its vendor CTS Corporation (Supply Chain Digest, 2010).

Mattle the maker of popular Barbie dolls and Hot Wheels cars, recalled nearly one million toy world wide in August 2007, because the toys manufactured during a two month period by its contracted vendors in China contained impermissible levels of lead (Story, 2007). Although, these vendors had been working with Mattle for years, their improper sourcing decision which allegedly caused these lead tainted toys had severe consequence for the entire supply chain. These cases highlight that supply chain performance is susceptible to variation for the constituent members.

As manufacturing firms outsource more parts and services to focus on their own core competencies, they increasingly expect their suppliers to deliver innovative and quality products on time and at competitive cost. When an underperforming supplier provides an innovative product or process technology (that may be sustainable long-term advantage to the buyer), the buyer may wish to protect this potential advantage and bring the work in-house by acquiring the supplier. But a better option for the buyer may be supplier development (Handfield et. al, 2000).

According to Burns and Whittle, (1995), companies must analyze their situation to determine if supplier development is warranted. To this end, Handfiled *et al*. (2000), classified commodities according to the company's relative importance of purchases using Kraljic portfolio approach. Kraljic's, (1983), approach includes the construction of a portfolio matrix that classifies products on the basis of two dimensions: profit impact and supply risk ("low" and "high"). The result is a 2x2 matrix and a classification in four categories: bottleneck, non-critical, leverage, and strategic items. Handfield *et al*, (2000) found that commodities in strategic supplies category are considered strategically important, difficult to substitute or purchase from alternative suppliers, and purchased in relatively high volumes therefore the suppliers of these commodities become the target for supplier development program.

There may be multiple suppliers of strategic supplies. Several studies have been focused on Supplier Selection Problem (SSP) that consists of analyzing and measuring the performance of a set of suppliers in order to rank and select them to improve the competitiveness of the entire supply system. Many conflicting factors should be taken into account in the analysis, both qualitative and quantitative. The most utilized methodology is represented by the well-known Analytical Hierarchical Process (AHP) (Saaty, 1980 and 1994) with its different variants. Basic versions of AHP are still widely used in the literature to deal with the SSP.

The evolution of supply relationships underlines that suppliers are required to have an adequate set of competencies to be part of a supply system capable of facing market competition (Esposito and Passaro, 2009). To this aim, customer firms have performed various actions and strategies: in particular the assessment processes has assumed a crucial importance. It represents a compulsory and critical starting point for the achievement of a collaborative customer–supplier system (de Boer *et al.*, 2001). Ranking is crucial but it does not solve the problem of identifying suppliers for development. Ranking suppliers from best to worst gives the idea of performance and scope for future improvement through benchmarks. It is not helpful in identifying suppliers with potential for development. To warrant supplier development, we need to identify the underperforming suppliers and involve them in development program to optimize their performance in order to amplify the performance of supply chain. In this paper at first we will demonstrate how AHP can be used by a company to rank its suppliers and subsequently using Pareto analysis to identify critical underperforming suppliers delivering strategic supplies to be included in supplier development program.

2. ANALYTIC HIERARCHY PROCESS

AHP can be used in making decisions that are complex, unstructured, and contain multiple attributes (Partovi, 1994). The decisions that are described by these criteria do not fit in a linear framework; they contain both physical and psychological elements (Mian and Dai, 1999). AHP provides a method to connect that can quantify the subjective judgment of the decision maker in a way that can be measured. In applying AHP to benchmarking, Partovi, (1994) describes the process in three broad steps: the description of a complex decision problem as a hierarchy, the prioritization procedure, and the calculation of results. AHP is a method of breaking down a complex, unstructured situation into its components parts, arranging these parts or judgments on the relative importance of each variable and synthesizing the judgments to determine which variables have the highest priority and should be acted upon to influence the outcome of the situation (Saaty, 1990). A problem is put into a hierarchical structure with the level-I reflecting the overall goal or focus of the decision (Saaty, 1990). The prioritization process is accomplished by assigning a number from a scale developed by Saaty, (1990) to represent the importance of the criteria. A matrix with pair wise comparisons of these attributes provides the means for calculation.

2.1 Method Application and Result

Step 1: Define the evaluative criteria used to select the optimal distribution network

Possible criteria that a company might consider when selecting a supplier are product cost, quality, service, and delivery. The mathematical process used in AHP is to first establish preferences for each supplier for each criterion. In this study the critical supplies of an automotive company were considered with six suppliers who supply these critical supplies.

Step 2: Establish each factor of the pair-wise comparison matrix

In this step, preferences between alternatives are determined by making pairwise comparisons. A judgment matrix is formed and used for computing the priorities of the corresponding elements. First, a criterion is compared pair-wise with respect to the goal. The judgment matrix, denoted as A, will be formed using the comparison. Let $A_l, A_2, \ldots, A_n$, be the set of stimuli. The quantified judgments on pairs of stimuli A_i, A_j, are represented by

$$A = [a_{ij}],\ i, j = 1, 2 \ldots, n \qquad (1)$$

The comparison of any two criteria C_i and C_j with respect to the goal is made using the questions of the type: of the two criteria C_i and C_j which is more important and how much. Saaty (1980) suggests the use of a 9-point scale to transform the verbal judgments into numerical quantities representing the values of a_{ij}. Larger number assigned to the pair-wise comparisons means larger differences between criteria levels. The entries a_{ij} are governed by the following rules:

$$a_{ij} > 0,\ a_{ji} = 1/a_{ij},\ a_{ii} = 1 \text{ for all } i \qquad (2)$$

This scale can be applied with ease to criteria that can be defined numerically as well as to those cannot be defined numerically. Relative importance scale is presented. The decision maker is supposed to specify their judgments of the relative importance of each contribution of criteria towards achieving the overall goal.

Step 3: Calculate the eigenvalue and eigenvector

Having recorded the numerical judgments a_{ij} in the matrix A, the problem now is to recover the numerical weights ($W_1, W_2, \ldots, W_n$) of the alternatives from this matrix. In order to do so, consider the following equation:

$$\begin{bmatrix} a_{11} & a_{12} \dots & a_{1n} \\ a_{21} & a_{22} \dots & a_{2n} \\ . & . \dots & . \\ . & . \dots & . \\ a_{n1} & a_{n2} \dots & a_{nn} \end{bmatrix} \cong \begin{bmatrix} W_1/W_1 & W_1/W_2 \dots & W_1/W_n \\ W_1/W_1 & W_1/W_2 \dots & W_1/W_n \\ . & . \dots & . \\ . & . \dots & . \\ W_n/W_1 & W_n/W_2 \dots & W_n/W_n \end{bmatrix} \quad (3)$$

Moreover, we multiply both matrices in Eq. (3) on the right with the weights vector $W = (W_1, W_2, \dots, W_n)$, where W is a column vector. The result of the multiplication of the matrix of pair-wise ratios with W is nW, hence it follows:

$$AW = nW \quad (4)$$

This is a system of homogenous linear equations. It has a non-trivial solution if and only if the determinant of $A - nI$ vanishes, that is, n is an eigenvalue of A. I is an $n \times n$ identity matrix. Saaty's method computes W as the principal right eigenvector of the matrix A; that is,

$$AW = \lambda_{max} W, \quad (5)$$

where λ_{max} is the principal eigenvalue of the matrix A. If matrix A is a positive reciprocal one then $\lambda_{max} \geq n$, [11]. If A is a consistency matrix, eigenvector X can be calculated by

$$A - (\lambda_{max} I)X = 0. \quad (6)$$

Here, using the comparison matrix, the eigenvectors were calculated by Eqs. (5) and (6).

Step 4: Perform the consistency test

The eigenvector method yields a natural measure of consistency. Saaty, (1990) defined the consistency index (CI) as

$$CI = \lambda_{max} - n/(n-1) \quad (7)$$

Where λ_{max} is the maximum eigenvalue, and n is the number of factors in the judgment matrix. Accordingly, Saaty (1990) defined the consistency ratio (CR) as

$$CR = CI/RI, \quad (8)$$

for each size of matrix n, random matrices were generated and their mean CI value, called the random index (RI). Where RI represents the average consistency index over numerous random entries of same order reciprocal matrices. The consistency ratio CR is a measure of how a given matrix compares to a purely random matrix in terms of their consistency indices. A value of the consistency ratio $CR \leq 0.1$ is considered acceptable. Larger values of CR require the decision-maker to revise his judgments.

Results of the consistency test and the CR of the comparison matrix from the available interview and previous data are all ≤ 0.1, indicating 'consistency''.

Step 5: Calculate the overall level hierarchy weight to rank the suppliers

The composite priorities of the alternatives are then determined by aggregating the weights throughout the hierarchy. The composite priorities of the alternatives are shown in table 1. Table 1 gives the serial ranking of all the six suppliers supplying critical supplies to the automotive company with their ideals, normals, and raw weight values. Based on the four criteria and scores in the table 1, supplier 1 is operating optimally whereas the rest are not.

Table 1. The overall synthesized priorities for the suppliers

Ranking Category		Ideals	Normals	Raw
1	Supplier 1	1.0000	0.2134	0.0813
2	Supplier 2	0.8820	0.1882	0.0717
3	Supplier 3	0.7448	0.1589	0.0605
4	Supplier 4	0.7425	0.1585	0.0603
5	Supplier 5	0.6927	0.1478	0.0563
6	Supplier 6	0.6224	0.1328	0.0506

2.1.1 Sensitivity Analysis

The final priorities of the alternatives are highly dependent on the weights attached to the main criteria. Small changes in the relative weights can therefore cause major changes of the final ranking. Since these weights are usually based on highly subjective judgments, the stability of the ranking under varying criteria weights has to be tested. For this purpose, sensitivity analysis can be performed based on scenarios that reflect alternative future developments or different views on the relative importance of the criteria. Through increasing or decreasing the weight of individual criteria, the resulting changes of the priorities and the ranking of the alternatives can be observed. Sensitivity analysis therefore provides information on the stability of the ranking. If the ranking is highly sensitive to small changes in the criteria weights, a careful review of the weights is recommended. Also, additional decision criteria should be included as a highly sensitive ranking point to a weak discrimination potential of the present set of criteria. For this purpose, the weights of the important criteria are separately altered, simulating weights between 0% and 100% (note that the weights of the other criteria change accordingly, reflecting the relative nature of the weights, i.e., the total weights has to add up to 100% in this paper). Sensitivity analyses are necessary because changing the importance of criteria requires various levels of cost, quality, service, and delivery w.r.t selecting the optimal performing supplier which can be prospective candidates for supplier developments.

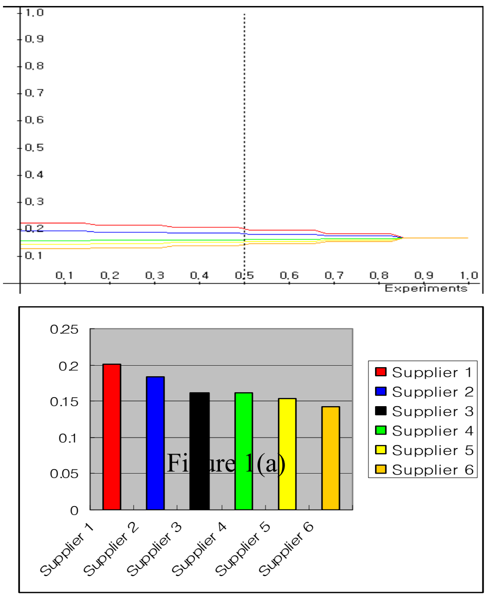


Figure 1(a)

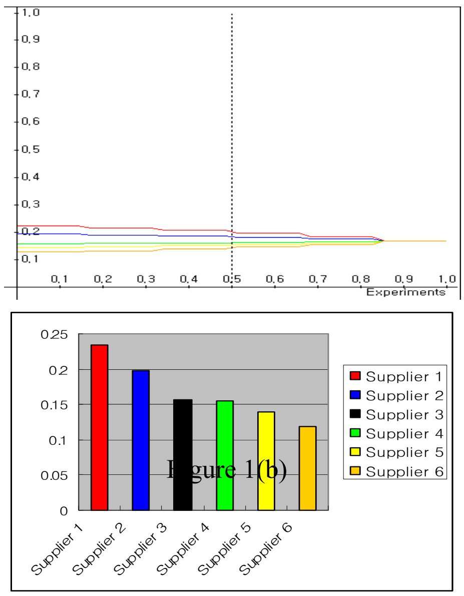


Figure 1(b)

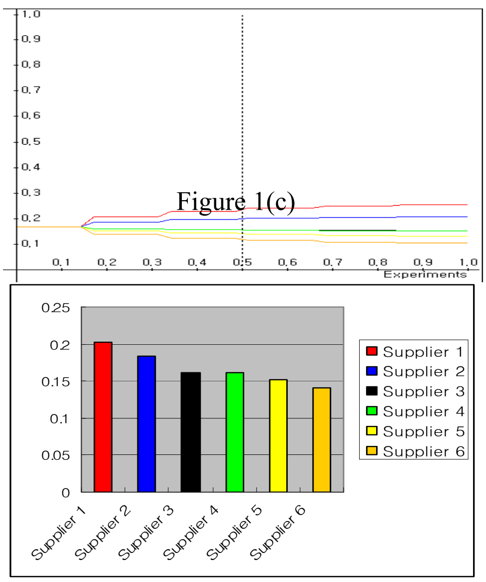


Figure 1(c)

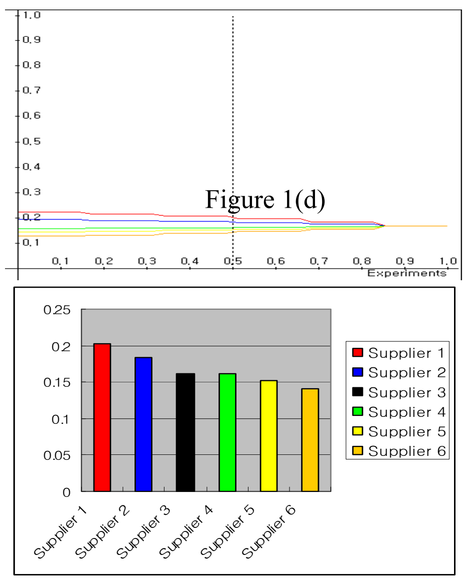


Figure 1(d)

Figure 1. The sensitivity graph of activities. (a) With independent variable cost (b) with independent variable quality (c) with independent variable delivery (d) with independent variable service

The sensitivity graph (figure 1) displays how the alternatives perform with respect to the change in scenario of all parameters. Performance sensitivity of alternatives has been analyzed when activity (independent variable cost) is increased by 50%, plotted on the X-axis and the priorities of the alternatives plotted on the Y-axis. Similarly for independent variables quality, service, and delivery, it is increased by 50%. Increasing the activity cost by 50%, as in figure 1 (a), there is a negligible change in the global weights (table 1). Same is the case with activity quality (figure 1 (b)) and service (figure 1(b)), however for activity delivery in figure 1(c) an increase by 50% tends to change the ideal values of suppliers 1, 2, 3, 4, 5, and 6 from its global values of (0.213, 0.188, 0.158, 0.158, 0.147, 0.132) to (0.234, 0.198, 0.156. 0.155, 0.140, 0.118). That means at point activities = 0.5, the score of supplier 6 decreases and supplier 2 increases. Before that sensitivity scores are consistent with the global scores.

2.2 Pareto Analysis

The Pareto analysis is useful in identifying suppliers for potential for development, as well as those that are underperforming low volume suppliers. From this Pareto analysis we find that 20% of suppliers are responsible for 80% of the poor performance of the chain as shown in figure 2. Supplier 1 is already above the cutoff point so there is no need for supplier development. However, supplier 2, 3, and 4 are target suppliers who are underperforming in the company's performance objective of cost, quality, service and delivery. Since these three suppliers contribute to the maximum percentage of underperformance, supplier development is therefore warranted.

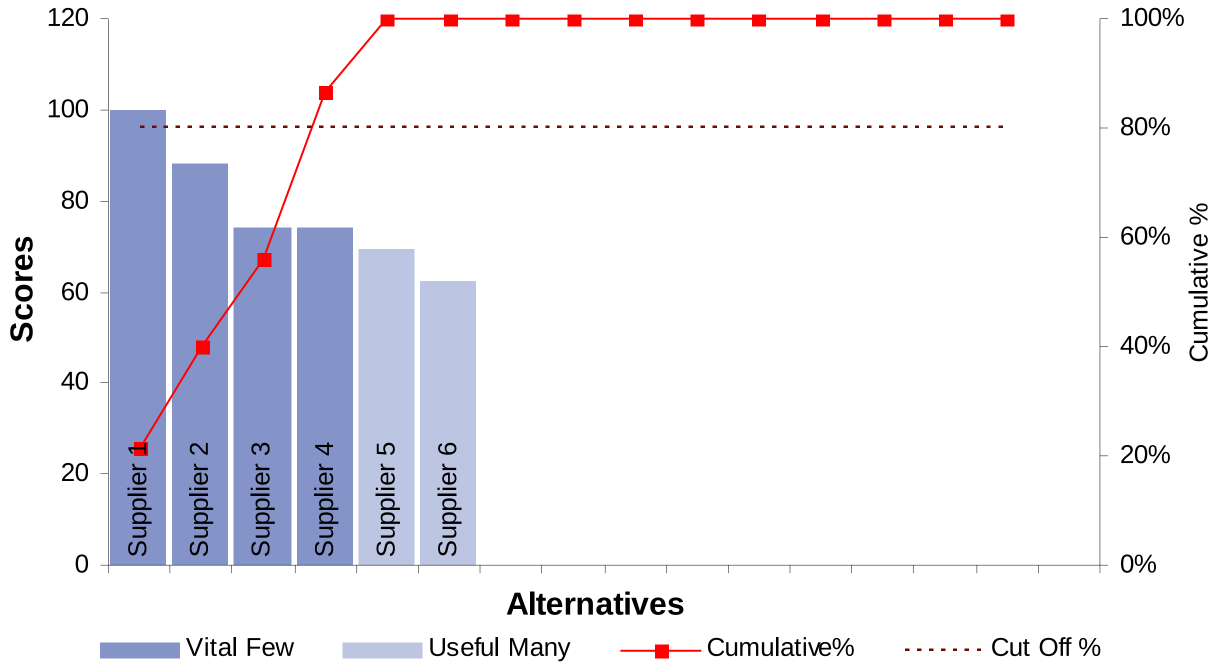


Figure 2. Pareto analysis of supplier performance

3. CONCLUSION

In this research we found that 20% of suppliers are responsible for 80% of the poor performance of the supply chain. Ranking is crucial but it does not solve the problem of identifying suppliers for development. Using AHP we ranked the suppliers from best to worst. This gives the idea of performance and scope for future improvement through benchmarks. The question was which suppliers to involve in supplier development. It is not possible to include all the underperforming suppliers into the supplier development program. Therefore the Pareto analysis gave us the result of vital few that should be included in the supplier development program. Since supplier 1 is performing beyond the cutoff point, it does not warrant buyer's development. Suppliers 2, 3, and 4 are the prospective candidates for supplier development since Pareto analysis shows they contribute the maximum to the supply chain performance. Suppliers 5 and 6 are useful many and can be eliminated from being involved in the program.

4. REFERENCES

Burnes, B., and Whittle, P. 1995. Supplier Development: Getting Started, *Logistics Focus,* 3: 10-14

de Boer, L., Labro, E., & Morlacchi, P. 2001. A review of methods supporting supplier selection, *European Journal of Purchasing & Supply Management*, 7: 75-89.

Esposito, E.., and Passaro, R. 2009. The evolution of supply chain relationships: An interpretative framework based on the Italian inter-industry experience. *Journal of Purchasing and Supply Management*, 15: 114-126.

Handfield, R.B., Krause, D.R., Scannell, T.V., and Monczka, R.M. (2000). Avoid the Pitfalls in Supplier Development, Sloan Management Review, 41: 37-49.

Kraljic, P. 1983. Purchasing Must Become Supply Management, *Harvard Business Review*, 61: 109-117.

Mian, S.A. and Dai, C.X., (1999). Decision making over the project life cycle: an analytical hierarchy approach, *Journal of Project Management*, 30: 40-52

Partovi, F.Y. 1994. Determining what to benchmark: an analytic hierarchy process approach, *International Journal of Operational Production Management*, 14: 25-39

Saaty, T.L. 1980. *The Analytic Hierarchy Process*. McGraw Hill International, New York.

Saaty, T.L. 1990. *Decision Making for Leaders*, RWS Publications, PA,

Saaty, T.L. 1994. Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process. RWS Publications, Pittsburg, PA.

Story L (2007) Lead paint prompts Mattel to recall 967,000 toys. New York Times, Aug 2

Supply Chain Digest Homepage. http://www.scdigest.com

Session B2: Design & Operation 1

·**Day1: Sep. 15 (Wed.)**

·**Time: 13:20 - 14:40**

·**Chair: Loo Hay Lee**

·**Room: Azalea & Lilac, 5F**

IDENTIFICATION OF ECONOMICALLY PROMISING PORT ARCHITECTURES VIA ENUMERATION AND CAPACITY EVALUATION

Jong Hoe Kim[1] and James R. Morrison[2*]

Department of Industrial and Systems Engineering
KAIST,
335 Gwahangno, Yuseong-gu,
Daejeon 305-701, Korea
e-mail:{windf , james.morrison}@kaist.ac.kr

Abstract: We develop a method to identify economically promising port architectures. We first enumerate feasible port architectures using a tool which generates feasible architectures for complex systems. We identify basic functions of ports and find options to satisfy each function. We enumerate feasible architectures by selecting options logically. As a result, 9 and 128 feasible architectures of traditional ports and offshore service system are generated, respectively. To evaluate and compare the architectures we develop rough-cut equations for ship staying time, throughput and cost (under deterministic assumptions). Such throughput based approximations serve to pare down the architectures to a promising handful. From there we conduct detailed simulations to compare them. As a result, we find which type of mobile harbor is cost effective and which container handling resource is more efficient at the land. We compare offshore service system with traditional ports for construction cost and throughput.

1. INTRODUCTION

Due to the historical and anticipated growth in container transport volume along with ever increasing competition (Korea Container Terminal Authority, 2006), container ports worldwide continuously strive to improve their capacity and service quality. Traditional solutions for capacity increase include optimization methods and the development of faster crane and yard systems. An alternate approach is to conduct offshore operations and to develop faster and economically superior interfaces between the offshore operations and the land berth.

A port is a complex system, in that there are many components and relationships between them. As such, there are many possible options for port expansion decisions, such as the type of berth, the type of crane and type of transfer unit. Therefore, one challenge to selecting good capacity improvement options is the vast array of potential solutions. The goal of this work is to develop a method to identify economically promising port architectures.

We first develop a representation of port systems based on the various functions they perform, such as unloading ships, porting containers across the land-sea interface and storing containers. For each such function, alternative solution concepts are identified. By selecting one solution for each function, possible port architectures may be obtained. Examples of port architectures include the traditional port (which uses quay cranes to provide the ship unloading functions) and the mobile harbor (which uses an offshore crane to unload the ships and a container transfer barge to port the containers to the land based port). We formalize the functions of the architectures and the solutions for each as an Object Process Network (OPN). OPN is a tool to generate feasible architectures for complex systems; it allows one to exclude architectures that are not logically possible. OPN software is then used to obtain a collection of 137 feasible port architectures, including 9 traditional and 128 offshore port concepts.

To compare the collection of architectures generated, we use rough cut capacity analysis to determine the number of components, such as cranes, berths and yard trucks, required to provide container ship service for a given market condition. The analysis also uses service time approximations to ensure that there are sufficient resources to serve ships within twenty four hours. We refer to an architecture with a specific number of units for each component as a configuration. We determine the construction cost of a configuration using published and projected cost figures; the operation cost is assumed to be a known percent of the construction cost. Based on these analyses, our approach enables the determination of promising architectures and the required number of components to provide an average service level. We study several test markets to determine which architectures are most promising.

2. ENUMERATION OF PORT ARCHITECTURE

To enumerate feasible architectures for a port, Object Process Network (OPN) was used. OPN (Koo, 2005) is a domain neutral, executable meta-language. This tool assists in scientific decision making for complex architectures and has been successfully adopted to generate complex system architectures (Simmons at el., 2005 and Rozenblum, 2007). It automatically generates feasible architectures by using a large number of possible options and relationships between the options. Since this tool exhaustively generates all architectural options, combinations of components that were not considered before will be produced. Thus, the process can lead to unexpected insights.

2.1 Process of enumeration

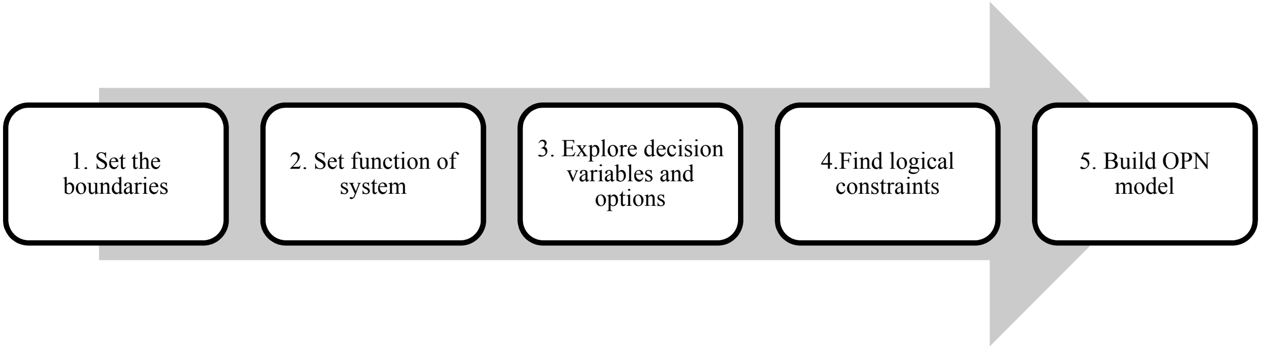


Figure 1. Process of generating of architectures

Figure 1 shows the process of generating architectures. In the first step, the boundaries of the system were selected as "container transportation system between containerships and the inland yard" for the main focus. Next, we identified a set of functions that the selected system from step 1 must provide. To provide a function, several decisions must be made. These decision variables each have several options. For example, the resource for unloading can be cranes, railed RORO, and wheeled RORO. After this process, logical constraints were built between options in order to prevent the generation of illogical architectures. With the output of process 3 and 4, the OPN model was built and used to generate feasible architectures.

2.2 Set the Boundaries

In order to focus on the most important objectives for the system and to consume a reasonable level of time and resources, we defined and set the highest specifications of the system. In our study, the boundary was a "container transportation system between containerships and the inland yard". We consider yard as the sink and source of the containers. Also, feeder ships were not considered in this system.

2.3 Select functions of the system

Before selecting the functions of the system, the system was classified into two modes, which were traditional mode and mobile mode. Traditional mode refers to traditional ports in which ships dock to the inland berth and unload cargo between the ship and berth by using quay cranes. Mobile mode refers to the mobile harbor concept in which container ships stay at sea and a system of container transport ships and offshore cranes moves the containers from the containership to the inland yard. Table 1 shows the functions of the traditional mode and the mobile mode.

Table 1. Functions of traditional mode and mobile mode

Traditional mode	Mobile mode
1. Moor /Stabilize	1. Moor /Stabilize (Ship & unloading unit)
2. (Un)load	2. (Un)load (Ship & unloading unit)
3. Berth	3. Transfer ((un)loading unit & travel unit)
4. Transfer(berth-yard)	4. Travel
5. Store	5. Moor/Stabilize (travel unit-Berth)
	6. (Un)load (travel unit -Berth)
	7. Form/unstack batch (at front berth->transfer unit)
	8. Berth
	9. Transfer
	10. Store

2.4 Explore decision variables and options of each function

There were several decisions to be determined for each function. What kinds of resource will be used for unloading at the berth? What type of transfer unit will be used? What type of berth will be used? Decision variables are the variables that characterize the architecture. Also, the option is the possible values for the decision variable. A set of seven decision variables was selected for the traditional mode and 15 decision variables were selected for the mobile mode. These decision variables also have a set of options. Also, only one option can be selected for one decision.

2.5 Logical constraints

There are logical constraints between the decision variables. These constraints prevent the generation of illogical architectures. All architectures must satisfy these constraints. Also, there are two kinds of logical constraints, which are first order constraint and second order constraint. First order constraints are between the two decision variables and second order constraints are between three variables. An example of a second order constraint is that the, mobile harbor cannot contain these three options for decision variables (unit type = non intergrated , unloading resource location = on mobile harbor and unload resource at berth = crane) in an architecture because the non-integrated type does not have a crane. There are two second order constraints for the traditional mode and 17 first order constraints and 8 second order constraints for the mobile mode.

2.6 Build OPN model to generate feasible architectures

Feasible architectures were generated via the OPN software with the selected decision variables and logical constraints. OPN-IDE [0.1.14] software and an Intel(R) Core(TM)2 Quad CPU Q6600 @2.40Hz, 2.00GB RAM computer were used. The OPN model of traditional mode and mobile mode produces 9 and 128 feasible architectures respectively out of a total combinatorial space that consists of 16 and 13824 combinations respectively. Figure 2 shows the OPN model for the mobile modes (left) and traditional modes (right). Table 2 shows several feasible architectures of the mobile mode. In this table, several well-known concepts are matched with generated architectures such as A1 type (Suh, 2008), mid stream operation (Wang, 1998), Hybrid Mobile Floating Port (HMFP) (Morrison et al., 2008) and B type (Yoon, 2008).

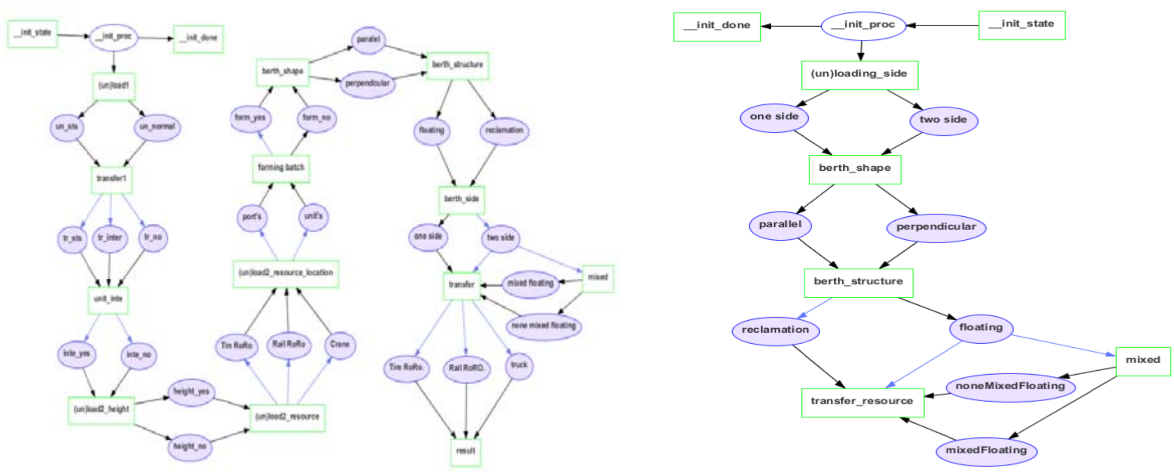


Figure 2 . OPN model of mobile mode (left) and traditional mode (right)

Table 2. subset of feasible architecture of mobile mode

Ship unload resource	Transfer Resource to TU	Unit integration	height adjust	Berth Unload resource	Unload Resource location	Forming Batch	Berth Shape	Berth Structure	Berth Side	Berth Mixed	Transfer To yard	Arch #	Remarks
roma	no	yes	yes	RailRoRO	ports	no	perpendicular	reclamation	oneSide	N/A	RailRoRO	3	A1 type
roma	no	yes	no	crane	units	no	parallel	floating	oneSide	N/A	truck	23	Mid stream operation
roma	interface	no	yes	RailRoRO	ports	yes	parallel	reclamation	oneSide	N/A	truck	70	HMFP
roma	interface	no	yes	TireRoRo	ports	yes	parallel	reclamation	oneSide	N/A	truck	99	E type option B
STS	STS	no	yes	TireRoRo	ports	yes	parallel	reclamation	oneSide	N/A	truck	109	B type option A

3. CAPACITY EVALUATION

The architectures were compared based on three metrics: throughput, cost, and ship staying time. A limit on the average containership staying time was set and the cheapest configuration of architectures meeting this constraint was found under the given market throughput. In order to do so, the cheapest configuration of the architecture for the given condition needed to be calculated. We made several assumptions to conduct these evaluations.

The interarrival times between ships and the cargo volume per ship are assumed deterministic and constant. Mooring speed is constant, irrespective of the amount of cargo and the architecture (20 min). The land yard is considered as the sink and source of containers (it has a large enough crane to absorb/release at 60TEU/hour). Due to safety issues of container stacks during transfer, mobile container transfer units (TU) moor at the berth perpendicularly when using the Roll on Roll off (RORO) system. TU moors parallel to the land berth when using a crane system because we assumed the crane spreader has standard shape. There is only one type of berth (one sideed berth constructed via reclamation). TU/Truck/RORO conducts single cycle operation (when these units travel between place A and B, these units travel with cargo when it travels from A to B, but these units are empty when they return from B to A). Ships receive service immediately upon entering the port. All ships have the same size and number of docking points. Since we use only three metrics, some architectures are dominated by other architectures. Therefore, we can narrow our focus to six operation modes for the mobile mode. Figure 3 shows the six operation modes of the mobile mode.

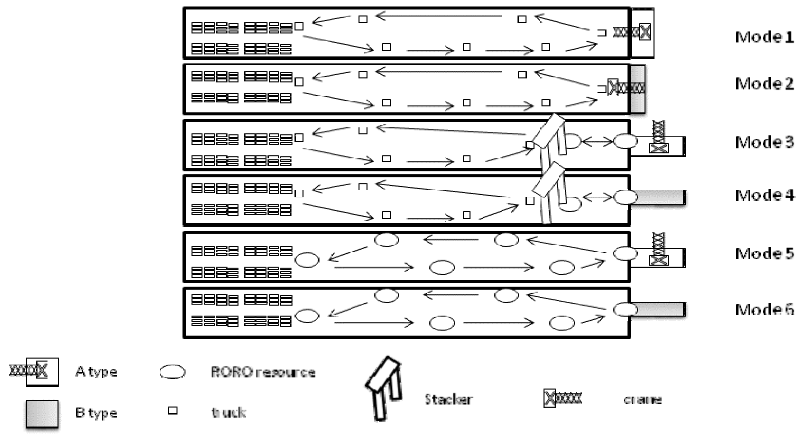


Figure 3. Six operation mode of mobile mode

In Mode 1, integrated (A type) TU units which have their own crane are used to transport cargo between the berth and ship. The crane on the unit is used to unload at the berth and ship. Trucks are used to transport containers from the apron (immediately after unloading) to the yard. In Mode 2, floating cranes that are set up at sea unload the container between the ship and non-integrated (B type) TU units that do not have a crane. B type units travel between the floating cranes and the land berths. The cranes that are set up at the berth unload the container between TUs and trucks. Trucks are used to transport containers from the apron (immediately after unloading) to the yard. In Mode 3, A type units are used to transport cargo between the berth and ship. The crane on the unit is used to unload at the ship. At the berth, RORO resources are used to transport the container between the TUs and stacker. The stackers transfer the containers between the RORO resources and trucks. Trucks are used to transport the containers' to the yard. In Mode 4, floating cranes setup at sea unload the containers between the ships and B type units. B type units travel between the floating crane and berth. At the berth, RORO resources are used to transport the container between the TUs and stacker. The stackers transfer the container between the RORO resources and trucks. Trucks are used to transport containers from the stacker to the yard. In Mode 5, A type units are used to transport cargo between the berth and ship. The crane on the unit is used to unload at the ship. At the berth, RORO resources are used to transport the container directly between the units and yard. In Mode 6, floating cranes setup at sea unload the containers between the ships and B type units. B type units travel between the floating cranes and the berth. At the berth, RORO resources are used to transport the container directly between the units and the yard.

3.1 Ship Service Time (SST) approximation for a single ship

In this section, we develop an equation that allows us to approximate the ship staying time. If every part of the system operates independently, every resource can work without waiting for other resource. In this situation, the cargo flow rate of the system is determined by the cargo flow rate of the bottleneck. The ship service time can thus be a approximated by $\mathrm{MAX}(\frac{\text{cargo volume}}{\text{cargo flow rate[i]}})$, where i varies over the container flows in the system. There are three container flows in the mobile harbor mode: between the ship and TU at sea, between the TU at sea and TU at berth, and between the TU at berth and the berth. We use the following notation: cargo is the container volume/ship, nD is number of docking points/container ship, nTU is the number of TU/container ship, nB is the number of berths/container ship, nTruck is the number of truck/berth, nRORO is the number of RORO resources/berth or unit, sShip is the TEU service rate at the ship/crane, sBerth is the TEU service rate/berth, cTU is maximum storage/TU, tShip is unloading time of TU at ship, tBerth is unloading time of TU at berth, tTra is travel time of TU between operation place and berth. At the berth, the cargo flow rate is calculated as MIN(nB,nTU) * sBerth. Between the berth and container ship, the max cargo flow rate is (cTU/(tShip+2*tTra+tBerth))*nTU. One TU can transfer cTU containers in one cycle time. For unloading on the ocean, the max cargo flow rate is min(nD,nTU)*sShip. Therefore, the SST can be approximated as

$$\mathrm{SST} \approx \max\left(\frac{\mathrm{cargo}}{\min(\mathrm{nD},\mathrm{nTU})*\mathrm{sShip}}, \frac{\mathrm{cargo}}{\left(\frac{\mathrm{cTU}}{\mathrm{tShip}+2*\mathrm{tTra}+\mathrm{tBerth}}*\mathrm{nTU}\right)}, \frac{\mathrm{cargo}}{\min(\mathrm{nB},\mathrm{nTU})*\mathrm{sBerth}}\right).$$

(1)

Approximation (1) is obtained by assuming independent operation between resources. However, in actuality, there is a relationship between two resources that meet each other. For example, if the number of customers is greater than the number of servers, customers often wait. There are two servers for TUs in port systems, the berths and ships. Waiting time depends on the entering pattern of TUs on the servers. In this paper, we assume all customers come to the server simultaneously. Also, the bunches of customers come to the servers after finishing processing the previous customers. Figure 4 shows the assumed arrival pattern.

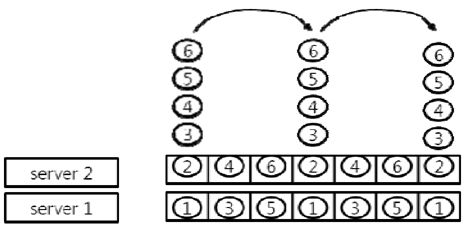


Figure 4. Assumed arrival pattern of customers on severs

In this case, the average waiting time for TUs can be calculated as

$$\mathrm{T_w} = \frac{\mathrm{N_S}*\mathrm{T}*\sum_{1}^{\left\lfloor\frac{\mathrm{Nc}}{\mathrm{Ns}}\right\rfloor-1} n + \left\{\left(\mathrm{Nc}-\left\lfloor\frac{\mathrm{Nc}}{\mathrm{Ns}}\right\rfloor*\mathrm{Ns}\right)*\left\lfloor\frac{\mathrm{Nc}}{\mathrm{Ns}}\right\rfloor*\mathrm{T}\right\}}{\mathrm{Nc}},$$

(2)

where Nc is the number of customers, Ns is number of servers, T is the service time. Therefore, the container flow of TU should be changed as

$$\frac{\mathrm{cargo}}{\left(\frac{\mathrm{cTU}}{\mathrm{tShip}+2*\mathrm{tTra}+\mathrm{tBerth}+\mathrm{Twb}+\mathrm{Tws}}*\mathrm{nTU}\right)},$$

(3)

where Twb is waiting time at berth (using equation for Tw and appropriate input values) and Tws is wating time at ship (using equation for Tw & appropriate input values)

3.2 Verification of rough cut equation

The rough-cut equation was compared with a time based simulator for the A type TU with a 240 TEU capacity. We tested the case in which nB < nTU. Since we assume deterministic ship arrivals and cargo volume, only one replication is needed for each case. We assumed 45 TEU/hour and 120TEU/hour for unload speed at ship and the berth. The number of TUs and berths varies from one to eight and the cargo volume/ship varies from 500TEU to 4500TEU. We compared 154 systems and the average absolute gap between the SST approximation and the simulation was is 6.49%.

3.3 Calculation of cheapest configuration under given market information

In the previous section, we calculated the SST of a single ship under a given number of resources and cargo volume/ship. In this section, we introduce a method to find the cheapest configuration of the architecture under a given market condition. From this, the cheapest architecture for a given market condition is obtained.

There is one constraint for the system, which is the service level for the ship. In this study, the level of service for the container ship was set to SST < 24 hours. The configurations of architectures that do not satisfy this constraint will be eliminated. If IAT is greater than SST, the port only needs one set of resources for a ship. However, if IAT is smaller than SST (that is, if an other ship enters the port before the previous ships), we need an extra set of resources to cover

another ship without a waiting time. SST/IAT is the average number of ships in port. For example, if SST is 24 hours and IAT is eight hours, then are three ships in the port in steady state. Therefore, we can estimate that the port needs [SST/IAT] sets of resources.
Algorithm to find cheapest configuration under given market information :

1. Get market information (IAT, cargo volume/ship),
2. Generate possible configuration combination,
(E.g. in mode 1, nB = 1~5, nTU = 1~10, nTruck = 10,20,30,40 => total of 200 combinations),
3. Filter configurations that do not satisfy 24 hour constraints,
4. Multiple [SST/IAT] to number of resource/ships ,
5. Find minimum cost configuration among feasible configurations.

4. RESULT

A variety of market conditions were studied by combining IAT and cargo volume from the options IAT = [1.5,2,2.4,3,4,6,12,18,24,30] hours and cargo = [500,900,1300,1700,2100,2500,2900,3300,3700,4100] TEU/ship. The smallest market size was 0.146 M TEU/year (IAT=30, cargo=500) and the biggest market size was = 2.4 M TEU/year (IAT = 1.5, cargo=4100). 240 TEU TUs were considered. Table 3 shows the range of input variables.

Table 3. The range of input variables for the resource (240 TEU TU)

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6
nFC N/	A	1~4	N/A 1~	4	N/A	1~4
nB	1~10	1~10	1~10 1	~10 1	~10	1~10
nTU	1~20	1~20	1~20 1	~20 1	~20	1~20
nTruck	10,20,30,40	10,20,30,40	10,20,30,40 10	,20,30,40 N/	A	N/A
nRORO	N/A	N/A	2~40 2	~40 2	~40	2~40

Figure 5 shows the cheapest configuration of Mode one for each market condition that meets the SST<24hours restriction.

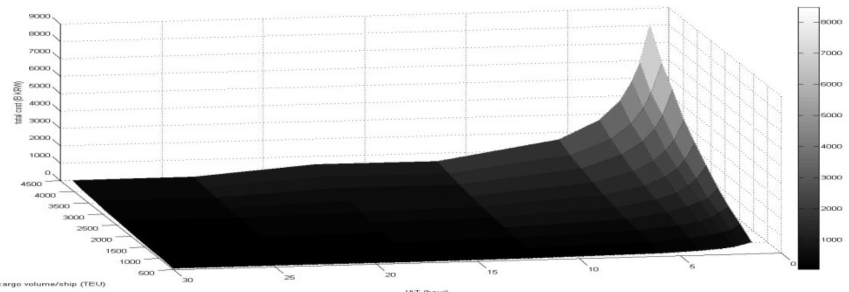


Figure 5 Cheapest configuration of Mode 1

From these graphs, it can be observed, and is not at all surprising, that the total cost increases as the cargo volume or IAT decreases. We can thus identify, the cheapest architecture for the given market condition. Table 4 shows the subset of cheapest configurations of each mode for a given market condition. The bold number represents the cheapest architecture for that market.

Table 4. The subset of cheapest configuration of each mode for given market condition

IAT	cargo	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6
1.5	500 6	84 5	79	497.64	492.88	464.76	**459.28**
1.5 9	00	1212	1040	822.6	875.14	823.52	**814.94**
1.5	1300 1	710 1	512	1210.22	1250.2	**1184.58**	1217.04
1.5	1700 2	280 1	904	1623.3	1635.66	**1544.1**	1602
1.5	2100 3	040 2	380	**1919.4**	2013.12 1	950.3 1	979.28
1.5	2500 3	840 2	848	2488.2	2478.84	2323.8	**2316.72**

1.5	2900 4	656 3	328	3087.36	2916.48	3136.2	**2698.88**
1.5	3300 5	856 3	808	3861.2	3304.6	3718.08	**3081**
1.5	3700 6	880 4	272	4387.2	3759	4211.84	**3507**
1.5	4100 8	480 4	992	5280.96	4067.2	5288.32	**3961.28**

4.1 Relationship between mobile Architectures: Integrated (A) and Separated (B)

The B type has the advantage of a faster crane located offshore and a faster TU travel speed. However, the unit cost of the floating crane is more expensive than the unit cost of the crane for the A type. It was observed that the A type becomes more favorable when the cargo volume is small and B is more favorable when the cargo volume is large. The cost gap between A and B grows as the cargo volume increases.

For the small size ship, the A type and B type need a small number of resources. A type automatically has four docking points on the ship, but the B type should have a floating crane (FC) offshore to provide docking points to TUs. Also, the unit cost of the floating crane was more expensive than the unit cost of a crane on a TU. If the B type has less FC located offshore than nTU, the waiting for TU will occur and it will result in increasing the SST and increasing SST/IAT, so that it will have more average ships in port. Therefore, the A type is preferred for the case of small cargo volume/ship. However, when the mobile harbor serves large cargo volume/ships, it needs many TUs and uses all the docking points of the ship. In this case, the total number of crane/ships is nTU for the A type and nD, which is four, for the B type. Therefore, B is preferred for the large cargo volume/ship.

4.2 Relationship between Crane and RORO system

Even though the RORO system has extra cost, such as link span, guided line, stacker, and pallet system, the berth speed is faster than the crane so it reduces the number of TUs and berths required. Also, the RORO system has a shorter berth length compared to cranes; this covers the extra cost of the RORO system. Therefore, in most cases, crane system has a more expensive total cost than the RORO system for a given market.

4.3 Relationship between RORO with stacker and RORO without stacker

Both kinds of architectures achieve similar berth speeds. The RORO with stacker system needs a fewer number of RORO resources because it divides the berth to yard transfer system into two parts. Therefore, the RORO system just handles half of the system. Also, the RORO system is not suitable for long distance container travel compared to the truck because of its speed. Therefore, it is more efficient for a truck to be in charge of long distance travel. However, in order to get this advantage, the RORO with stacker system needs an expensive stacker to divide the system into two parts. If we reduce the expensive stacker cost, a RORO with stacker system can be competitive.

4.4 Comparison with traditional port

A traditional port (Korea Container Terminal Authority, 2008) which has been built recently was examined. Based on information of construction cost, the mobile harbor concepts are always cheaper. However, the cost of the traditional port was based on a 25% crane utilization assumption. Therefore, if we cut the traditional costs by one fourth, the traditional port will be more cost effective. Note that throughout we have just considered the construction cost. However, unlike the Gwang Yang port (Korea Container Terminal Authority, 2008), Hong Kong and Singapore have very expensive land purchase cost and taxes. Therefore, they charge very expensive container handling fees. In those markets, the traditional port suffers from expensive land cost while the mobile harbor has less effect on these expensive costs because it uses a very short length of the seashore. Also, the A type is a competitive solution for a very small market. It avoids huge initial construction costs. In this case, only one or two A type TUs are required with a basic quay.

5. CONCLUSION

To generate mobile harbor architectures, an OPN model was developed. This model allowed us to generate feasible architectures by automatically enumerating all possible port design decisions. Logically infeasible architectures were pruned, so that 128 feasible architectures remained. In this paper, these feasible architecture were grouped together in an evaluation process. Only three criteria were used. The results demonstrated several dominance relationships between the architectures. For example, if we use a floating quay at the berth, it is more expensive but can be more rapidly deployed than the reclamation quay. Since construction time was not considered, reclamation was the most dominant option. In future detailed evaluation, we can use ungrouped feasible architectures with various metrics.

In the evaluation process, rough-cut equations (under deterministic assumptions) were developed based on throughput while approximate resource contention formulas determined the ship service time for a given configuration. These approximations were implemented in MATLAB code to study a wide range of port configurations to determine the cheapest cost solution for a given port. For each port (ie. total annual/TEU and TEU/ship), the architectures were compared and the best (cheapest) one was identified. From this result, several conclusions were drawn. The integrated type A MH units were generally good when the TEU/ship was less than 2500TEU/ship including import and export cargo. The separated type B MH units were good when TEU/ship was greater than 2500 TEU/ship including import and export cargo. The RORO systems at the berths for transfer of container from TU to land were superior to cranes for the same purpose. This is due to a dramatically faster speed/meter of berth in spite of an increased infrastructure cost/berth. RORO without stacker and direct transfer of container to yard is better than the intermediate stacker crane before the yard. This is due to a high cost of the stacker cranes. Also, this evaluation method can be used as a quick evaluation system.

The total cost of mobile harbors is not in favorable condition compared to the traditional port(Korea Container Terminal Authority, 2008). However, it is a viable solution for a very small market or a place in which the cost of land is very expensive, such as Hong Kong and Singapore. To reinforce the accuracy of this study, several studies are recommended. Since the construction time of a mobile harbor is half the traditional time, the NPV of each system should be compared. Also, this evaluation approach assumed a deterministic inter arrival time and cargo volume. Instead of using deterministic values, using random values will increase realism.

6. REFERENCE

B. H. Y. Koo (2005). A meta-language for systems architecting. *Ph.D. dissertation, Massachusetts Institute of Technology*, Cambridge, MA.

James Jixian Wang (1998). A container load center with a developing hinterland: a case study of Hong Kong. *Journal of Transport Geography*.

James R. Morrison, Taesik Lee, Jong Hoe Kim, Minsung Kim, Kyuhyeon Shin and Inkyung Sung (2008). Hybrid mobile floating port. *South Korean patent application*, Filing number: 10-2008-0121984, Filing date: December 3, 2008.

Korea Container Terminal Authority (2007). Analysis of trend of container transportation in 2006.

Korea Container Terminal Authority (2008). Information of Gwang Yang container terminal construction.

N. P. Suh (2008). Mobile Harbor to Improve Ocean Transportation System. *Korean patent application* No. 2008-41981, filed May 2008.

W. L. Simmons, B. H. Y. Koo, and E. F. Crawley (2005) Architecture generation for Moon-Mars exploration using an executable meta-language. *Proceedings of AIAA Space* 2005, 30 August - 1 September, Long Beach, CA, 2005.

W. L. Simmons, B. H. Y. Koo, and E. F. Crawley (2005). Space systems architecting using meta-languages. *56^{th} International Astronautical Congress.*

Yongsan Yoon (2008). Fast Container Transfer Module. South *Korean patent application*, Filing number: 10-2008-0090655, Filing date: September 16, 2008.

Z. Rozenblum (2007). Object-process networks & object-process diagrams - implementation issues for oil exploration systems. Master's thesis, Massachusetts Institute of Technology.

A SIMULATION FRAMEWORK FOR MEGA-SIZED CONTAINER TERMINALS

Sun Zhuo[1], Tan Kok Choon[2], Lee Loo Hay[3], Chew Ek Peng[4]

[1]Centre for Maritime Studies
National University of Singapore
12 Prince George's Park
Singapore 118411
e-mail: sunzhuo@nus.edu.sg

[2]Department of Decision Sciences, NUS Business School
National University of Singapore
15 Kent Ridge Drive
Singapore 119245
e-mail: biztkc@nus.edu.sg

[3]Department of Industrial and Systems Engineering
National University of Singapore
10 Kent Ridge Crescent
Singapore 119260
e-mail: iseleelh@nus.edu.sg

[4]Department of Industrial and Systems Engineering
National University of Singapore
10 Kent Ridge Crescent
Singapore 119260
e-mail: isecep@nus.edu.sg

Abstract: This study aims to develop a simulation framework to facilitate the design and evaluation of layouts of mega-size container terminals with multiple yards and berths which may not be regular shaped. Most studies related to container terminals focus on regular shapes of yards and berths. However, due to pre-existing geographical conditions, many container terminals in the world take on various layouts for their multiple yards and berths. The proposed framework in this study uses a GIS to help terminal layout generation. A generalized container terminal operations simulation model is then applied. The results from the simulation model can be used to evaluate overall efficiency of each layout design and thereby help port designers make future plans that suit the given geographical conditions.

1. INTRODUCTION

International trades between countries have been rapidly increasing in volume and value with globalization. Majority of the cargo and goods that have to be moved in international trades are packed into standardized containers for better management and most cost-effective transportation. As important maritime transportation nodes, major container terminals are getting more and more congested as the volume of containers passing through them increased. Container terminal operators typically arrange for various container vessels to berth at a quay and use terminal equipments (quay cranes, yard cranes and yard trucks) to lift, transport and store containers in the yards and then transport them to some hinterland or other vessels later. Computer simulation is usually the preferred approach to evaluate the efficiency of container terminals in meeting future demands or in new configurations. Most container terminal simulation models tend to assume that berths and yards are regularly shaped (straight linear berths and rectangular yards) and simplify their topographical relationships (one berth is associated with one yard). However, due to pre-existing geographical conditions, many container terminals in the world take on various irregular shapes and layouts for their multiple yards and berths. By considering these real geographical constraints, building simulation models becomes much more challenging and requires more effort.

Container terminal simulation models consist of resources and logics. Resources accordingly represent physical entities in the real world, such as gates, berths, yards, vessels, quay cranes, yard cranes, yard trucks and containers. Logics imitate the behaviors of those movable resources (terminal equipments) and decision processes made by control centers of container terminals. The conventional approach of building a simulation model for a typical container terminal is as follows. First, based on the berth length and yard length and width, construct a model of the berth and yard. Then add quay cranes, yard cranes and yard trucks as movable resources and write logics for them. Finally, write logics for decision processes as external controllers of resources. When running the simulation, vessels will be randomly generated as the testing demand. Figure 1 shows a typical container terminal in a simulation. Its layout has been simplified to concentrate on container operations. Because of the pre-existing geographical conditions the real layout of a container terminal may not be regular shaped as shown in Figure 2. Although a big terminal with multiple berths and yards can be decomposed and transformed to a few regular-shaped terminals, the operations of various equipments in the terminal cannot be decomposed simply. In fact, as experienced terminal managers would point out, the terminal layout may impact significantly the operations of terminal equipments. Thus, when building simulation models for these container terminals more things need to be considered. This study introduces a simulation framework to facilitate building this type of simulation. A Geographic Information System (GIS) has been used to help build the container terminal layout. Based on the layout, terminal resources will be allocated by a sharing policy. Then a generalized container terminal operations simulation model is applied. The results from the simulation model can be used to evaluate overall efficiency of each layout design and thereby help port designers make future plans that suit the given geographical conditions.

In the rest of the paper, a literature review of previous related works is given in Section 2, followed by problem definition in Section 3. The design of the simulation framework is presented in Section 4 while a case study is discussed in Section 5. Finally, conclusions are given in Section 6.

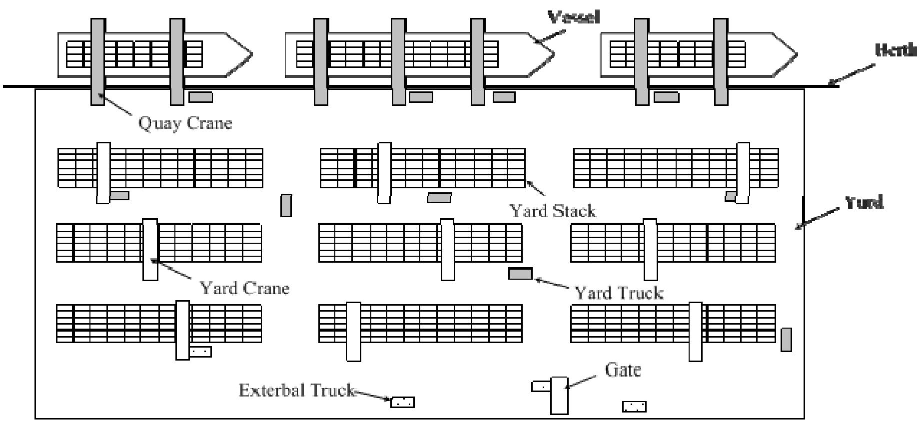


Figure 1. A Typical Container Terminal in a Simulation

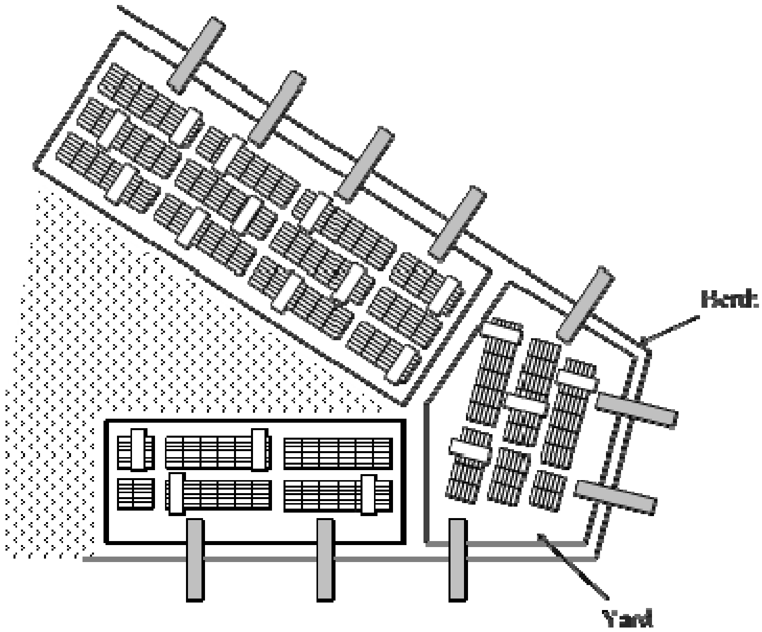


Figure 2. The Real Layout of a Container Terminal

2. LITERATURE REVIEW

In the recent decade, because of advancements in computer science, simulation, especially discrete-event simulation, has been widely used for investigating container terminals.

Gambardella et al. (1998) built a decision support system focused on the resource allocation problem and used a terminal simulation model as a test bed for checking the validity and robustness of the system. Nevins et al. (1998a, 1998b) developed a seaport simulation model that computes throughput capability and determines resource utilization at a high level of details. The simulation allows for multiple cargo types as well as multiple ship types. Shabayek and Yeung (2002) proposed a simulation model of Kwai Chung Container Terminals in Hong Kong to investigate the extent by which a simulation model could be used to predict the actual container terminal operations with a higher order of accuracy. Demirci (2003) used simulation experiments to find that the most critical bottlenecks were created by loading/unloading vehicles and an investment strategy was applied to the model for load balancing in the port. Parola and Sciomachen (2005) tried to simulate the logistic chain of the northwestern Italian port system as a whole to evaluate a possible future growth of container flow. Bielli et al. (2006) elaborated on an object-oriented design of a simulator for container terminals. Every piece of equipment, queue, or area is implemented as an object and communications among objects are implemented as messages. Zeng and Yang (2009) proposed a simulation model for scheduling loading operations in container terminals and then embedded this model in an optimization routine to find an optimal scheduling scheme.

There are some papers related to simulation and container terminal layout. Nam et al. (2002) used a simulation model with four scenarios to examine the optimal size for the Gamman Container Terminal in terms of berths and quay cranes. Liu et al. (2004) used simulation models to demonstrate the impact of automation and terminal layout on terminal performance. In particular, two terminals with different but commonly used yard configurations were considered for automation using AGVs. Lee et al. (2007, 2008) developed simulation models to investigate how different vehicles and different yard layouts can affect the efficiency of port operations. They further built a program named Automated Layout Generation to generate different simulation models. Hartmann (2004) introduced an approach for generating scenarios of sea port container terminals. The scenarios can be used as input data for simulation models. Furthermore, they can be employed as test data for algorithms to solve optimization problems in container terminal logistics such as berth planning and crane scheduling. Ottjes et al. (2006) introduced a generic simulation model structure for the design and evaluation of multi-terminal systems. This model is constructed by combining three basic functions: transport, transfer, and stacking. Petering (2007) developed a comprehensive simulation model to address issues in terminal design, storage and retrieval location and yard crane control. We note that none of above works in the literature is concerned with arbitrarily shaped container terminals with multiple berths and yards.

3. PROBLEM DEFINITION

This study is motivated by the problem currently faced by terminal designers in the port of Singapore, the world's largest and busiest container port. A big part of the port currently occupies an area that is very close to the central business district. As the price of land around the port rises with the demand for commercial land, it becomes inevitable that the container terminals have to make way. The obvious choice is to relocate the container terminals to the south western corner of the island and the only way is to reclaim land from the sea there for container terminal operations use. The designers need to decide on the sizes and shapes of the reclaimed land to construct a new mega-sized container terminal that will fit into the given geological constraints. It becomes necessary that the mega-sized container will include multiple berths and yards which will make things more complicated. Terminal designers thus face the challenge in analyzing and evaluating different designs.

This study covers the areas of Computer-Aided Design, Port Operations and Simulation. In computer-aided design system, accurate geographic information needs to be used. Port operations should take into account the problems associated with resource sharing. The simulation structure needs to be flexible and reusable for automated model generation.

4. FRAMEWORK DESIGN

The proposed framework is designed based on the MicroCity framework (http://microcity.sourceforge.net). There are three software layers designed in the proposed framework, namely; function layer, application layer and user layer. In the function layer, many base functions are implemented to support higher layers. In the application layer, the terminal layout will be analyzed and a set of generic container terminal operations will be used to control the terminal. Then a multi-agent system will be built to simulate the processes in the container terminal. In the user layer, users can input parameters and planning templates through the specific graphical user interface. A geographical information system in this layer can help the user to design the terminal layout. These layers are elaborated in the following subsections.

4.1 Function Layer

The function layer in the proposed framework is implemented in C/C++ to provide fast functions to higher layers. Geographical functions include coordinate transformations, line and polygon analyses and manipulations, shape overlap analyses and etc. For building simulation in higher layers there are two series of functions: random number generators and discrete event scheduler. A MersenneTwister random number generator (Matsumoto and Nishimura, 1998) has been embedded to facilitate generating random numbers in different distributions. The discrete event scheduler is implemented by using a priority queue and Lua(http://www.lua.org) coroutine to facilitate building discrete event simulations in the application layer.

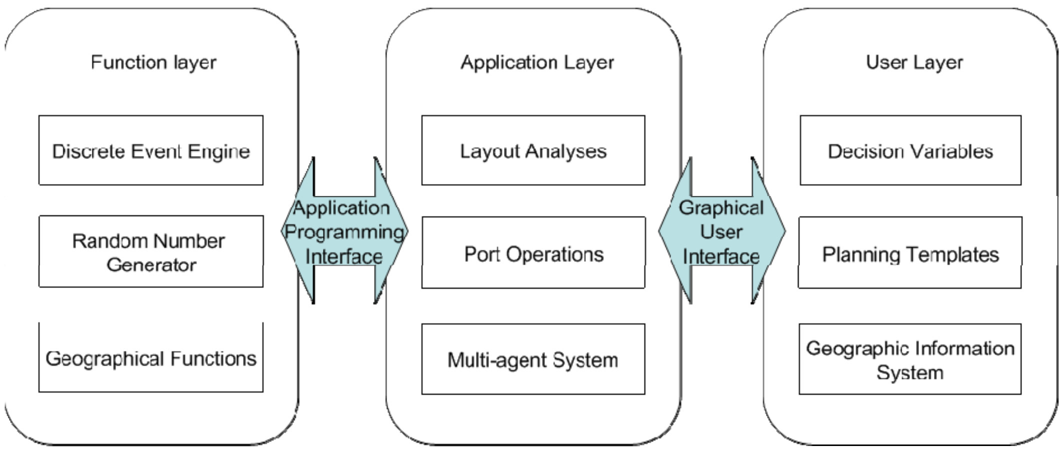


Figure 3. Three Software Layers in the Proposed Framework

4.2 Application Layer

Container terminal layout is analyzed in application layer with geographical functions mentioned in section 4.1. The geographical information of berths and yards can be extracted respectively. Container stacks are generated for every yard based on the yard shape and settings. Then a certain number of yard cranes are assigned to each stacking zone in a yard according to the ratio of the number of stacks in the zone to the total number of stacks in the terminal. Quay cranes are evenly distributed along the berths.

When the layout analysis is done, port operation policies need to be set for future use in the simulation. There are nine generic operation algorithms built into the framework, namely Berth Allocation, Quay Crane Assignment, Quay Crane Scheduling, Storage and Retrieval Location Assignment, Yard Crane Deployment, Yard Crane Dispatching, Yard Truck Dispatching, Yard Truck Routing and Yard Sharing. Most of them consist of simple rules like first-in-first-out, free-rotating, free-random and etc which can be extended by users with external modules. Here two algorithms need to be highlighted: berth allocation and yard sharing. To minimize the inter-yard and inter-berth container flow vessels carrying same groups of containers are likely to be allocated in a same berth. This study assumes this problem has already been solved. So that vessels are separately generated for every berth. Than the normal berth allocation can be applied to allocate vessels at certain positions in the continuous berth. There is another issue related to irregular container terminal with multiple berths and yards. If the number of berths and yards are equal and every berth has its dedicated yard then the whole terminal can be decomposed to several independent parts. If this situation does not happen, yards can be shared by berths. Figure 4 illustrates the yard sharing policy. First, buffers are created for every berth. Every yard stack covered by a buffer can be accessed by the corresponding berth. If a yard stack is covered by two or more buffers this stack is shared by several berths. The first container arriving at this stack will lock the stack to this container's source berth until the stack become empty.

To generate simulation model this study uses multi-agent system implemented in the application layer of the framework. Multi-agent systems can effectively simulate complex systems like container terminals. Every agent are implemented as a coroutine in the multi-agent system and represented as an entity in the real world. Decision processes are implemented as normal functions rather than coroutines for central control purpose. For instance, there are dozens of yard trucks traveling in the yard and interacting with other equipments. The jobs sequence in each yard truck is scheduled by the yard truck scheduler in the control center of the terminal. To model this sub system every yard truck can be represented as an agent coroutine and the yard truck scheduler can be treated as a central controling procedure. The yard truck scheduler will regularly pause all of the agent coroutines to make plans for yard trucks. Then the simulation will be resumed and all of the agent coroutines are activated simultaneous to continue the yard truck behaviors.

4.3 User Layer

The user layer of the framework is built upon the graphic user interface of MicroCity. Users can not only set parameters save and load planning templates through the user layer but also design accurate layout of terminals with a geographic information system. Satellite maps can be imported as the background to calibrate the layout design. The geographic information will be used in the application layer afterwards. The simulation animation and results can also be displayed in GIS to let users check the real-time simulation status.

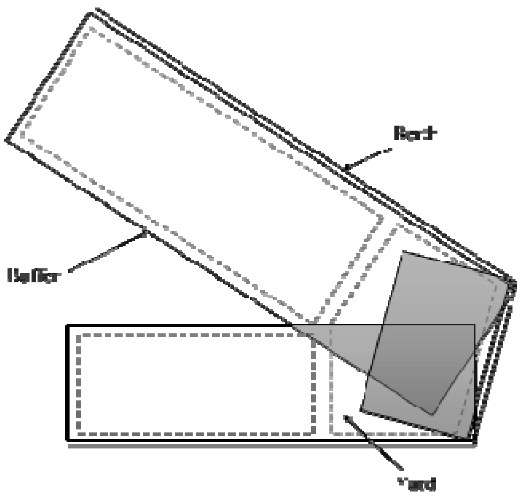


Figure 4. The Yard Sharing Policy

5. CASE STUDY

To test the practicability of this framework the Brani Terminal of Singapore is chosen as an example. Figure 5 shows the design phase a container terminal. A satellite map is imported as background to help design. Blue lines in this figure represent the berths. Red polygons represent yards. Because of the strict geographical conditions, the layout of the whole terminal is irregular.

Figure 6 is a screenshot of a running simulation applied in the same terminal. The terminal resources are proper allocated according to the layout. Vessels are generated randomly to test the terminal's efficiency. An experiment is performed to show how the number of yard cranes in this terminal can affect the terminal productivity which is measured by quay crane GCR (Figure 7). As the number of yard cranes increases the quay crane GCR first steeply increases then slows down. If a container terminal has limited resources these experiments can help the terminal operator to put reasonable invest on equipments.

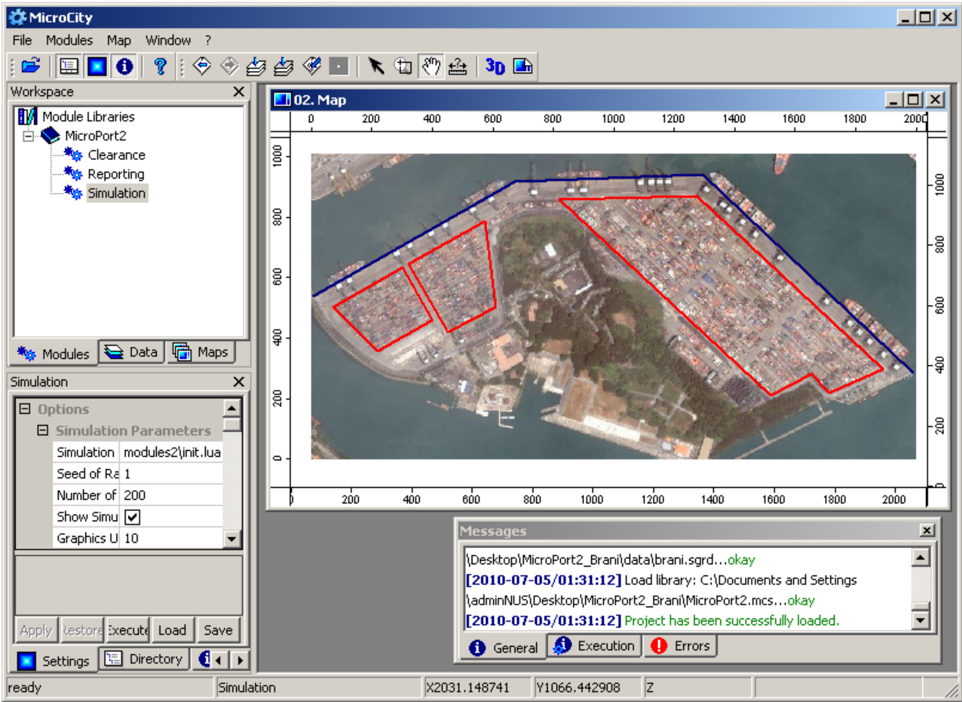


Figure 5. The Layout Design Phase of a Container Terminal

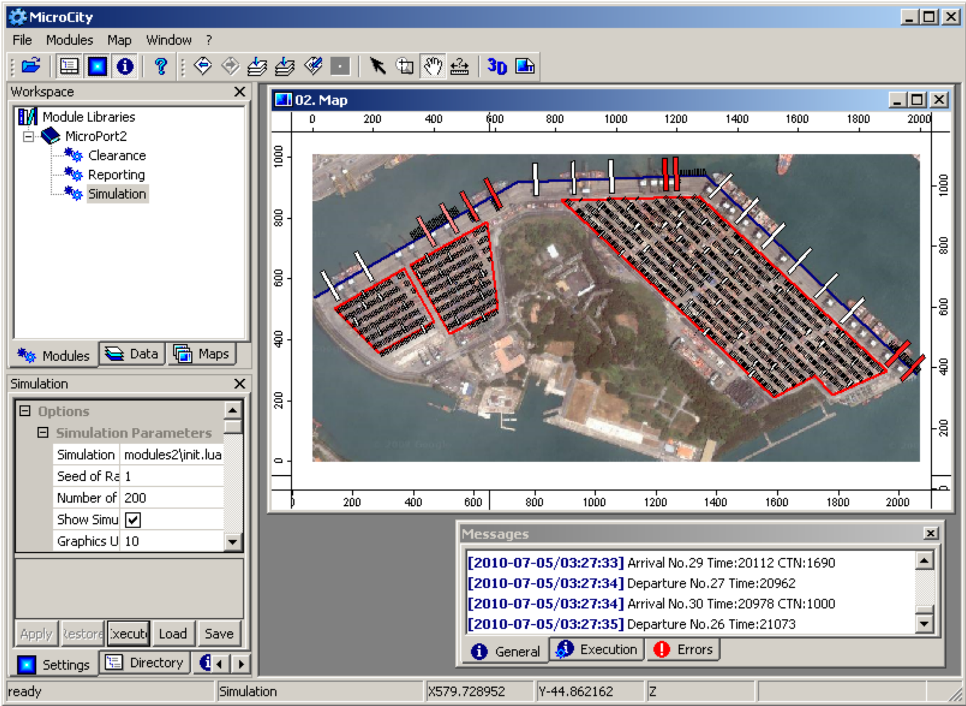


Figure 6. The Simulation Phase of a Container Terminal

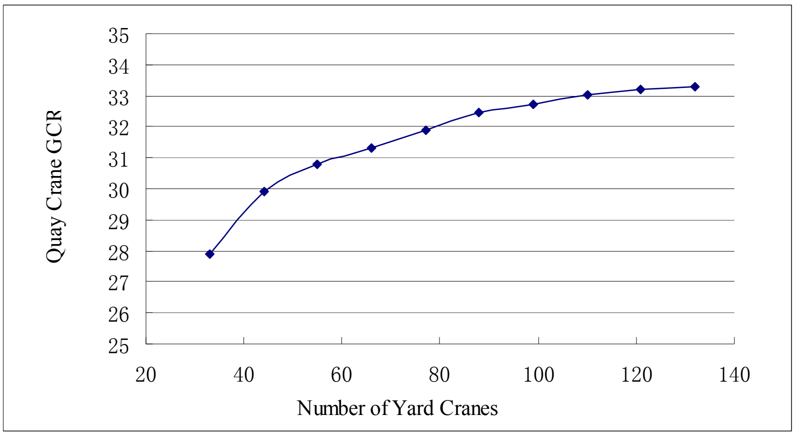


Figure 7. Quay Crane GCR for Various Number of Yard Cranes

6. CONCLUSIONS

This study introduced a general simulation framework for mega-sized container terminals. These terminals have multiple berths and yards which are likely to be not regularly shaped. The three software layers in the proposed framework provide convenient features in designing and developing simulation models for these container terminals.

The function layer provided basic functions for the application layer. The application layer analyses the terminal layout and provides nine generic decision processes for operations. Simulations are built through a multi-agent system implemented in this layer. The user layer provides friendly GUIs to facilitate inputting planning and design parameters. The GIS in the user layer can assist users to visually and accurately design terminal layout.

From the case study we can see that a comprehensive simulation model for a real container terminal can be generated on this framework. Various analyses can be performed by using the framework to provide useful information for port operators and designers. Obviously, the framework will greatly reduce the modeling effort and contribute to the integrated efficiency solutions for container terminals.

7. REFERENCES

Bielli, M., Boulmakoul, A. & Rida, M., 2006. Object oriented model for container terminal distributed simulation. European Journal of Operational Research, 175(3), 1731-1751.

Demirci, E., 2003. Simulation Modelling and Analysis of a Port Investment. SIMULATION, 79(2), 94-105.

Gambardella, L.M., Rizzoli, A.E. & Zaffalon, M., 1998. Simulation and Planning of an Intermodal Container Terminal. SIMULATION, 71(2), 107-116.

Hartmann, S., 2004. Generating scenarios for simulation and optimization of container terminal logistics. OR Spectrum, 26(2), 171-192.

Lee, L.H. et al., 2007. A simulation study on the uses of shuttle carriers in the container yard. In Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come. pp. 1994–2002.

Lee, L.H. et al., 2008. A study on port design automation concept. In Proceedings of the 40th Conference on Winter Simulation. pp. 2726–2731.

Liu, C. et al., 2004. Automated guided vehicle system for two container yard layouts. Transportation Research Part C: Emerging Technologies, 12(5), 349-368.

Lua: The Programming Language Lua. http://www.lua.org

Matsumoto, M. & Nishimura, T., 1998. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Transactions on Modeling and Computer Simulation (TOMACS), 8(1), 3–30.

MicroCity: A Spatial Analysis and Simulation Framework. http://microcity.sourceforge.net

Nam, K., Kwak, K. & Yu, M., 2002. Simulation Study of Container Terminal Performance. Journal of Waterway, Port, Coastal, and Ocean Engineering, 128(3), 126-132.

Nevins, M.R., Macal, C.M. & Joines, J.C., 1998a. A Discrete-Event Simulation Model for Seaport Operations. SIMULATION, 70(4), 213-223.

Nevins, M.R., Macal, C.M., Love, R.J. et al., 1998b. Simulation, Animation and Visualization of Seaport Operations. SIMULATION, 71(2), 96-106.

Ottjes, J. et al., 2006. Simulation of a multiterminal system for container handling. OR Spectrum, 28(4), 447-468.

Parola, F. & Sciomachen, A., 2005. Intermodal container flows in a port system network:: Analysis of possible growths via simulation models. International Journal of Production Economics, 97(1), 75-88.

Petering, M., 2007. Design, analysis, and real-time control of seaport container transshipment terminals. University of Michigan.

Shabayek, A.A. & Yeung, W.W., 2002. A simulation model for the Kwai Chung container terminals in Hong Kong. European Journal of Operational Research, 140(1), 1-11.

Zeng, Q. & Yang, Z., 2009. Integrating simulation and optimization to schedule loading operations in container terminals. Computers & Operations Research, 36(6), 1935-1944.

AN OPTIMAL ALGORITHM FOR SINGLE YARD CRANE SCHEDULING

Byung-Hyun Ha[1*] and Byung-Cheon Choi[2]

[1]Department of Industrial Engineering
Pusan National University
San 30, Jangjeon-dong, Geumjeong-gu
Busan, Republic of Korea
e-mail: bhha@pusan.ac.kr

[2] Department of Business Administration,
Chungnam National University
79 Daehangno, Yuseong-gu
Daejeon, Republic of Korea
e-mail: byungcheon.choi@gmail.com

Abstract: In container terminals, one of the most important performance measures is the berthing time of container vessels. The competitive advantage of terminals comes from their ability to serve container transporters, such as container vessels and trailers, as quickly as possible, and the efficient operations of yard cranes are one of the key factors for the high level of service. In this paper, we consider the problem of scheduling single yard crane in a container block. The goal of our study is to get the yard crane schedule that minimizes the average waiting time of trailers, which is measured by the difference between the ready time and the start time of a job. We explorer the several properties in the problem and suggest an optimal approach based on sweeping concept. We validated the effectiveness of our idea using numerical experiments.

1. INTRODUCTION

A container terminal in a seaport is the interface between land and marine transportation of containerized cargo. It plays an important role in intermodal transportation by storing containers temporarily on a container yard, in order to resolve the indispensible discrepancies between schedules of various transport modes. In most terminals, containers are stored in the form of container blocks on a container yard. Yard cranes pile up containers, delivered by container trailers, into blocks, and retrieve containers from blocks and load them into trailers. In container terminals, one of the most important performance measures is the berthing time of container vessels. Hence, the competitive advantage of terminals comes from their ability to serve container transporters, such as container vessels and trailers, as quickly as possible, and the efficient operations of yard cranes are one of the key factors for the high level of service.

In this paper, we consider the problem of scheduling single yard crane in a container block. There are two types of jobs assigned to yard cranes: i) retrieving a container from a block and loading it to a trailer and ii) picking up a container from a trailer and storing it to a block. A container block consists of container bays and a trailer stops directly by the bay, where a container to be retrieved is placed or where one to be stored will be place. To handle a job, a yard crane first moves from its current position to the place where the trailer requesting the job arrives, and then it serves the trailer. The ready time of a job is defined as the time when the trailer with the job arrives by the bay where the trailer will be served, and the start time of a job as the time when a yard crane completes the movement to the bay. The goal of our study is to get the yard crane schedule that minimizes the average waiting time of trailers, which is measured by the difference between the ready time and the start time of a job. Similar to other single-machine scheduling problems, a schedule of the problem is defined by a sequence of jobs. The problem has been studied in the several literatures, including Kim, Lee, and Hwang (2003), Kim and Kim (1999), and Ng (2005), and they have majorly suggested the practical rules for the yard crane dispatching.

The organization of this paper is as follows. In Section 2, we review the related research and identify the various perspectives on the problem and the possible applications. In Section 3, we precisely define the problem and suggest a mixed-integer programming model. An optimization algorithm by the concept of sweeping is presented in Section 4, and the performance of the suggested algorithm is discussed in Section 5. Lastly, we draw the conclusions in Section 6.

2. LITERATURE REVIEW

The problem in this study is characterized as a single-machine problem with the ready time of a job, the same processing time regardless of a job, and the sequence dependent setup time, which is specifically determined by the bays where two consecutive jobs are related (Bianco, Dell'Olmo, and Giordani, 1999). From the results of the scheduling research, the problem with the general setting is classified as NP-hard, because its special case is the problem with the ready time of a job and the total completion time as the objective function ($1 \mid r_j \mid \Sigma C_j$), which has been proved as strongly NP-hard (Lenstra, Rinnooy Kan, and Brucker, 1977). When the ready time constraint is relaxed and the minimization of the maximum completion time of a job is considered as the objective function, the problem becomes polynomially solvable (Pinedo, 2008). Baptiste (2000) considered the identical parallel machine scheduling problem with the constant processing time and the minimization of the total cost function, with special structure, determined by the completion time of a job. He proposed a dynamic programming approach that considers every possible release time point and proved that the problem is polynomially solvable. Hence, when the ready time of a job is not considered in our problem, it also becomes polynomially solvable, because it is a special case of the problem in Baptiste (2000). When two consecutive jobs are at the same bay, the setup time (the moving time of a yard crane) is zero. Thus, the setup time can be considered as the family setup time by regarding the jobs at the same bay are in the same family, and our problem is the single machine problem with the ready time of a job and the family setup time (Baker and Trietsch, 2009, Allahverdi, Gupta, and Aldowaisan, 1999).

The problem in this study can also be interpreted as the travelling repairman problem (TRP). The Line-TRP is the TRP problem with the special distance structure, where the jobs are considered as located along a line. Without the release dates and with the zero processing times, Line-TRP can be solved in polynomial time of $O(n^2)$ (Afrati *et al.*, 1986) but it was mentioned as an open problem in Tsitsiklis (1992) when the problem is one with the general processing times. When the release dates are considered with the zero processing times, the problem is NP-hard but open whether it is strongly NP-hard or not (Sitters, 2004). With the general processing time, the problem is strongly NP-hard, because its special case is $1 \mid r_j \mid \Sigma C_j$ as well. When not only the release dates but also the deadlines are considered, the problem becomes even harder. In that case, some bounds can be imposed on the problem input to make the problem more tractable. When a bound B is imposed on the number of nodes, the problem with the release dates, the deadlines can be solvable in $O(B^2 n^B)$. However, when the general processing times are considered, it is also strongly NP-hard, because $1 \mid r_j \mid \Sigma C_j$ is reduced to the problem, again. When a bound D can be given on the number of the active jobs at every time point, where a job is active at time t if t is in the interval between the release date and the deadline of the job, the travelling salesman problem (TSP) with the release dates and the deadlines can be solved in time of $O(TD^2 2^D)$, where T is an upper bound on the time horizon of the problem, while TRP with the same settings is NP-hard even in the case of the zero processing times (Tsitsiklis, 1992).

3. PROBLEM DESCRIPTION

There are n jobs given with the ready time r_j of job j. We assume the common processing time of all jobs and it is denoted by p. Further, it is assumed that the moving speed of a yard crane is constant and the setup time is proportional to the distance of the bays where two consecutive jobs are located. Without the loss of the generality, we suppose that a yard crane moves one bay per one time unit. The location, where job j is realized, is denoted by b_j. Initially, a yard crane located at the location l_0. Let C_j and y_{ij}, respectively, be the completion time of job j and the temporal precedence between job i and j. A mixed-integer programming for the problem in this study is as follows with the decision variables y_{ij}'s, where M is a sufficiently large number.

$$\text{min.} \quad \sum_{j=1}^{n} \frac{C_j - p - r_j}{n} \tag{1}$$

$$\text{s.t.} \quad |l_0 - b_j| \le C_j - p \qquad \forall j \tag{2}$$

$$r_j + p \le C_j \qquad \forall j \tag{3}$$

$$C_i + |b_i - b_j| + p \le C_j + M(1 - y_{ij}) \qquad \forall j \tag{4}$$

$$C_j + |b_i - b_j| + p \le C_i + M y_{ij} \qquad \forall i\ \forall j(>i) \tag{5}$$

$$y_{ij} = 0 \text{ or } 1 \qquad \forall i\ \forall j(>i) \tag{6}$$

Note that the objective of minimizing average waiting times is equivalent to the one of minimizing total completion times, because

$$\text{min.} \sum_{j=1}^{n} \frac{C_j - p - r_j}{n} \equiv \text{min.} \sum_{j=1}^{n} C_j - p - r_j \equiv \text{min.} \sum_{j=1}^{n} C_j \tag{7}$$

We will use the above equivalence when we identify the dominance rule in Section 4.

As mentioned in the previous section, the Line-TRP with the release date s and the zero processing times is NP-hard. Hence, it is clear that the problem above is more difficult because it has the processing time with strictly greater than zero. Nevertheless, the special properties, compared to the Line-TRP, are in the problem. First, every setup time, represented as $|l_i - l_j|$ above, is an integral number, and it can be interpreted as the same distances between two adjacent nodes in the perspective of Line-TRP. Second, the common processing time p is relatively large compare to the moving time between bays. For example, p is usually from 20 to 30 time units in practice while a yard crane moves by 1 bay in 1 time unit. Hence, the processing time is possibly represented by a certain integer. In the following section, we exploit the properties and suggest an optimal algorithm.

4. AN OPTIMAL APPROACH

In this section, we first assume the common processing time to be zero, and develop an optimal algorithm. In the end of the section we provide the extension to the processing time greater than zero.

Define J as a set of all jobs. Let a state S be the 4-tuple (t, l, A, z) where t, l, A, and z denote a time point, a location of a yard crane, a set of jobs already handled, and the total completion time of the jobs in A, respectively. The interpretation of the state is that a yard crane is located on bay l at time t and it has already handled jobs in A with the total completion time z. By determining the next job to be scheduled, the state $S = (t, l, A, z)$ can be followed by the new state $S' = (t', l', A', z')$, such that

$$l' \in \{b_j | j \in J - A\} \tag{8}$$

$$t' = \max\{|l - l'|, \min_{j \in \{j | b_j = l'\}}\{r_j\}\} \tag{9}$$

$$A' = A \cup \{j | r_j \leq t', b_j = l'\} \tag{10}$$

$$z' = z + \sum_{j \in A' - A} t' \tag{11}$$

The basic idea in our optimal approach is to sweep every possible state in ascending order of t, from time zero. Note that, when the next states are derived using Equation (8), not all the possible locations need to be considered, because some states can be considered based on the new state when $t' = |l - l'|$ holds in Equation (9). One of our conjectures is that, from the properties discussed in the previous section, the number of states does not increase too highly, even though it is possibly very large, especially when the dominated states are pruned effectively. There can be many possible dominant rules between the states and we describe an effective rule used by our algorithm as follows: state $S = (t, l, A, z)$ is dominated by $S' = (t', l', A', z')$, when

$$A \subset A' \tag{12}$$

$$t - |l - l'| \geq t \tag{13}$$

$$z \geq z' - z^*(A' - A, t, l) \tag{14}$$

where $z^*(A' - A, t, l)$ is the minimum total completion time when a yard crane handles all of jobs in $(A' - A)$ by starting from bay l at time t, and there is an optimal algorithm for obtaining $z^*(X, t, l)$ with the time complexity with $O(n^2)$, when $r_j \leq t$ for all $j \in X$ (Afrati *et al.*, 1986). While running the algorithm, the states that have not been dominated by others are kept as the frontiers. When a state is derived, it is compared with the frontiers and is pruned if there exists a frontier state which dominates the newly derived state. Otherwise, the new state becomes a frontier, and then each former frontier is checked whether it is dominated by the new frontier and pruned.

The proposed sweeping approach is not difficult to be extended to the problem with the processing times greater than zero with the following properties. First, the schedules with the processing order agreeable to the ready times of jobs in a specific bay constitute a dominant set. Second, when a yard crane completes a job, if there are the waiting jobs, i.e., the jobs with the completion time not less than the ready time, in the same bay with the completed job, the waiting jobs should be scheduled next. Third, $z^*(A' - A, t, l)$ can be still polynomially solvable because

$$\sum_j C_j = \sum_k C_{[k]} = \sum_k \sum_{i=0}^{k-1} \left(s_{[i][i+1]} + kp\right) = \sum_k \sum_{i=0}^{k-1} s_{[i][i+1]} + \frac{n(n+1)p}{2} \tag{15}$$

where $[k]$ is the kth job in the sequence and s_{ij} is the setup time when job j is handled immediately after job i.

4. EXPERIMENTAL RESULTS

For the experiments, we use the input, n, $l_{\max}$, and γ, from which there are generated n jobs with b_j, uniformly distributed from 1 to $l_{\max}$, and r_j, uniformly distributed from 1 to $(n{\cdot}\gamma{\cdot}l_{\max})$. The experiments are carried out using a PC with Intel Core2 Duo 2.53GH, 4GB RAM, and Windows XP.

Table 1. Average Computation Times using ILOG CPLEX (in sec.)

		γ									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
	20	1.5 0.	5 0.	3 0.	1 0.	2 0.	1 0.	1 0.	1	< 0.0	< 0.0
	25	4.7 1.	1 0.	8 0.	5 0.	8 0.	6 0.	3 0.	6 0.	6 0.	1
	30	73.7	13.2 1.	3 0.	7 0.	6 0.	7 0.	7 0.	8 0.	7 0.	7
n	**35**			405.1	134.5	120.3 76	.0 72	.3 46	.4 33	.1 24	.8
	40			294.4 25	8.9 23	4.3 15	6.1 14	1.0 10	3.7	92.3	52.6
	45				285.4 24	5.1 19	9.6 21	1.2 12	0.8	68.1 12	3.4
	50				437.6 28	6.6 26	0.0 22	1.6 17	6.2 17	3.1 18	3.8

Table 2. Average Computation Times using the Proposed Algorithm (in sec.)

		γ									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
	200	68.0 2.	5 0.	9 0.	5 0.	3 0.	3 0.	2 0.	2 0.	2 0.	2
	250	141.1 3.	8 1.	4 0.	9 0.	6 0.	5 0.	4 0.	3 0.	3 0.	3
	300	239.7 8.	5 2.	4 1.	4 0.	9 0.	8 0.	6 0.	5 0.	5 0.	4
n	**350**	372.2	10.7 3.	8 2.	1 1.	5 1.	1 0.	9 0.	8 0.	7 0.	6
	400	561.1	16.7 5.	5 2.	8 1.	9 1.	5 1.	2 1.	0 1.	0 0.	9
	450	778.0	23.4 7.	1 4.	0 2.	5 2.	0 1.	6 1.	4 1.	3 1.	1
	500	1250.3	31.8 8.	5 4.	5 3.	2 2.	5 2.	1 1.	8 1.	6 1.	4

Table 1 shows the average computation times of solving the problem instances by the mathematical programming in Section 3 using ILOG CPLEX. The empty cell in the table denote the computation time greater than 1 hour. We used ten instances per setting. Our optimal algorithm is implemented using C++ language and the computation times for all the instances used for Table 1 are less than 1 sec. Table 2 shows the computation time of our algorithm for the large-size problem instances.

5. CONCLUSIONS

In this paper, we consider the problem of scheduling single yard crane in a container block. The goal of our study is to get the yard crane schedule that minimizes the average waiting time of trailers, which is measured by the difference between the ready time and the start time of a job. We explorer the several properties in the problem and suggest an optimal approach based on sweeping concept. We validated the effectiveness of our idea by numerical experiments.

Even though the problem presented in this study is NP-hard in general setting, it is still open whether the problem is polynomially solvable or not, when the properties suggested in Section 3 are considered. Hence, it would be challenging to study the time complexity of the problem given together with the properties.

Commonly, there are two different types of trailers that request operations of a yard crane: the external trailers from outside of a terminal and the yard trailers operating internally within a terminal. The former is in charge of the container transportation between container terminals and container shippers, and delivers containers through gates of terminals. The latter transfers containers between a yard and a berth, where quay cranes load and discharge containers for a vessel, or between container blocks in a yard. That is, it is in charge of the container flow within a container thermal. Hence, each job has its own weight inherently. The possible direction of further research lies not only in the relaxation of the common processing times assumption but also in including the weight of jobs.

ACKNOWLEDGEMENTS
This work was supported by the Grant of the Korean Ministry of Education, Science and Technology (The Regional Core Research Program/Institute of Logistics Information Technology).

6. REFERENCES

Afrati, F., Cosmadakis, S., Papadimitriou, C.H., Papageorgiou, G., and Papakostantinou, N. (1986) The Complexity of the Travelling Repairman Problem. *Theoretical Informatics and Applications, 20:* 79-87.

Allahverdi, A., Gupta, J.N.D., and Aldowaisan, T. (1999). A Review of Scheduling Research Involving Setup Considerations. *Omega, 27:* 219-239.

Baker, K.R. and Trietsch, D. (2009). *Principle of Sequencing and Scheduling*. John Wiley & Sons.

Baptiste, P. (2000). Scheduling Equal-length Jobs on Identical Parallel Machines. *Discrete Applied Mathematics, 103(1-3):* 21-32.

Bianco, L., Dell'Olmo, P., and Giordani, S. (1999). Minimizing Total Completion Time Subject to Release Dates and Sequence-dependent Processing Times, *Annals of Operations Research, 86:* 393-415.

Kim, K.H., Lee, K.M., and Hwang, H. (2003). Sequencing Delivery and Receiving Operations for Yard Cranes in Port Container Terminals. *International Journal of Production Economics, 84:* 283-292.

Kim, K.H. and Kim, K.Y. (1999). An Optimal Routing Algorithm for a Transfer Crane in Port Container Terminals. *Transportation Science, 33.*

Lenstra, J.K., Rinnooy Kan, A.H.G., and Brucker, P. (1977). Complexity of Machine Scheduling Problems. *Annals of Discrete Mathematics, 1:* 343–362.

Ng, W.C. (2005) Crane Scheduling in Container Yards with Inter-crane Interference. *European Journal of Operational Research, 164:* 64-78.

Pinedo, M.L. (2008). *Scheduling: Theory, Algorithms, and Systems, 3rd Ed.* Springer.

Sitters, R.A. (2004). *Complexity and Approximation in Routing and Scheduling*. Ph.D. Thesis, Technische Universiteit Eindhoven.

Tsitsiklis, J.N. (1992). Special Cases of Traveling Salesman and Repairman Problems with Time Windows. *Networks, 22:* 263-282.

AN EXPERT SYSTEM BASED ON MULTI-REASONING MECHANISM FOR PORT MACHINE DIAGNOSIS

Yi Ding *, GuoLong Lin, Yanping Meng

Logistics Research Center
Shanghai Maritime University
1550 Pudong Ave, ShangHai China
E-mail:dingyi1018@yahoo.com.cn

Abstract: Expert system plays an important role in port machine diagnosis, which aims at automatic equipment test for higher availability and efficiency of port operations. In this study, a port machine diagnosis expert system is proposed based on multi-reasoning mechanism. Relying on the knowledge acquired from the experienced experts in the port machine engineering, the system builds a library of relative experience and a set of rules of reasoning and estimating. Multi-reasoning mechanism that simulates the decision-making process of domain experts is employed to achieve reliable diagnosis results. The reasoning machine integrates artificial neural network, uncertain decision making and decision tree, which complements each other by sustainable growing voting mechanism. The effect of this multi-reasoning mechanism is evaluated and validated by means of Matthew's Correlation Coefficient. The system incorporating the mechanism is successfully designed, implemented and applied in Shanghai Port.

1. INTRODUCTION

The modern port operation is becoming more and more complicated and the collaboration of different type of port machine is required to be more reliable and secure. The interruption of operation caused by the unpredictable faults of the system will have huge economic loss and even heavy casualty as a consequence. On both the practice and theory level, it is of great importance to develop the practical research on port machine diagnosis in order to improve the accuracy and the sensitivity of diagnosis, to lower the rate of false alarm and to increase diagnosis efficiency.

The Diagnosis expert system is the application of expert system to the diagnosis Domain. It makes full use of such strong ability to process the knowledge of the experts as acquiring the knowledge, memorizing the knowledge and reasoning. Based on the characteristic signals (including the normal and abnormal ones) and other diagnostic information, the diagnosis expert system indentifies sub-system or relation of some specific level which causes faults and finds out initial reasons for these faults.

By researching and applying multi-reasoning mechanism, this paper mainly works out the basic theory and design of port machine diagnosis expert system. After studying the characteristics, advantages and disadvantages of several traditional reasoning mechanisms thoroughly, the artificial neural network reasoning、decision tree reasoning and uncertain decision reasoning are combined to do research on the multi-reasoning mechanism and based on which the Diagnosis Expert System reasoning mechanism is developed and the knowledge base is designed according to the requirement of Expert System and port diagnosis domain .

2. EXPERT KNOWLEDGE ACQUISITION AND REPRESENTATION

The port machine diagnosis expert knowledge in this system includes the expert personal knowledge and the domain knowledge. The personal knowledge refers to the personal experience and theories summarized from the long-time practice by the domain expert. It is reserved to the expert and differentiates the domain expert from other domain staff. The domain knowledge refers to the one known by the common domain staff.

In the process of system developing, the system research staff should read the technical materials on the port diagnosis and know the basic domain concepts, the difficulties and the way of thinking so that they can be understood when communicating with the domain staff. For these domain staff, it is not proper to have them know the implementation details of the system. Because they do not have the professional computer knowledge (such as computer language and program design), the domain staff uses natural language when building knowledge. For the convenience of knowledge entering by the experts, this system applies graphic way of entering knowledge based on the fault tree. According to the requirement of the system, a set of expert knowledge entering form is provided for the domain experts to fill in. Take table1 as an example.

Table 1. Signal parameters of diagnosis process

Machine type	diagnosis clause name	signal name	Signal number	Signal amplitude	Signal cycle
…					

In the part of the expert knowledge described by the rules, there are only two relationships in the diagnosis knowledge base from the view of logical knowledge representation: rule and fact. The facts included machine type, diagnosis clauses and signal codes, as shown in figure 1, which describes the specific requirements of port machine diagnosis. The rules which is the central part of diagnosis expert knowledge base, is exclusively appointed through port machine type, diagnosis clauses and signal codes. The diagnosis consequence is made by the signal codes and corresponding Status.

Besides gaining the knowledge from the original information provided by the experts, the system can also increase the accuracy of reasoning by such learning mechanism as artificial neural nets, uncertain deduction and decision tree and therefore better the knowledge base which also saves the data structure and weight values of these mechanisms.

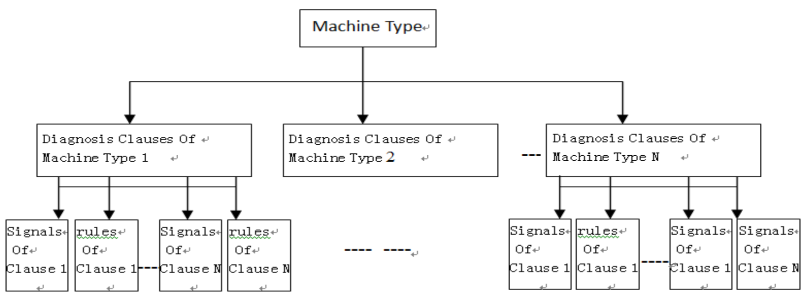


Figure 1. Knowledge base hierarchical structure

3. RESEARCH AND APPLICATION OF MULTI-REASONING MECHANISM

3.1 Artificial Neural

BP neural net is a Simplex Communication and multilayer forward artificial neural network. Besides the input and output layers, BP net has one or several hidden layers. There is no connection between the nodes on the same layer. Every node is a single nerve cell, of which transfer function is commonly sigmoid type function (Timo et al., 2003).

In the diagnosis expert system, the right result is reasoned by the well-trained neural networking, which is set for each diagnosis clauses in the knowledge base. The six signals of diagnosis clause 10 can be expressed as an input vector of six dimensions, each of which represent one signal, so the input layer is set to have 6 input neural nodes. The output layer is set to have 6 output neural nodes due to the number of diagnosis consequence. The next step is to determine the number of layers and nodes within of the hidden layer. In this system, according to the changes of node number of input and output layers in different diagnosis clauses, the node number of the hidden layer is decided as （int） Sqrt(number of input nodes*number of output nodes). Meanwhile, the rules in the clause 10 are transferred to the training samples in order to train the BP artificial neural networking.

A reasoning training module and a reasoning prediction module constitute the Artificial Neural Network module and are built on knowledge base system and port machine diagnosis platform separately as shown in figure 2.

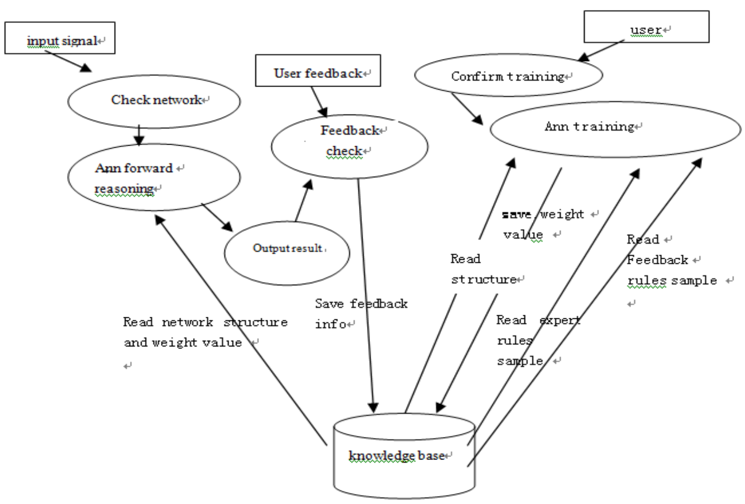


Figure 2. Artificial neural network reasoning flow diagram

3.2 Uncertain Decision Making

In this system, the results and application of traditional reasoning mechanism are restricted by the expert knowledge, which is limited in comparison with the real situation. The Fuzzy match uncertain reasoning will work as the major Uncertain Decision reasoning method (S.Horikawa, 2002). In this method, input signals and expert knowledge rule antecedent are matched to form the uncertain rules. The Fuzzy match uncertain reasoning is chosen to put high reliance on the expert knowledge, and therefore the uncertainty of the proof could be accurately solved. The continuous feedback reasoning trainings on Initial rules could help to increase the creditability of uncertain rules rapidly and accordingly come to a correct conclusion more easily.

The loss of the rule antecedent makes it a failure to determine the rule consequent. The aim of fuzzy match uncertain reasoning is to solve the signal loss, which will caused by the delay or blockage in the processing of the signal transfer; and the incompleteness of the rule base makes it difficult to reason correctly. The diagnosis task can be completed by uncertainly reasoning the all previous real diagnostic data. What's more, the system, by analyzing all previous diagnostic data, can come to an answer to adjust the credit of uncertain rules and the flow diagram is shown in figure 3.

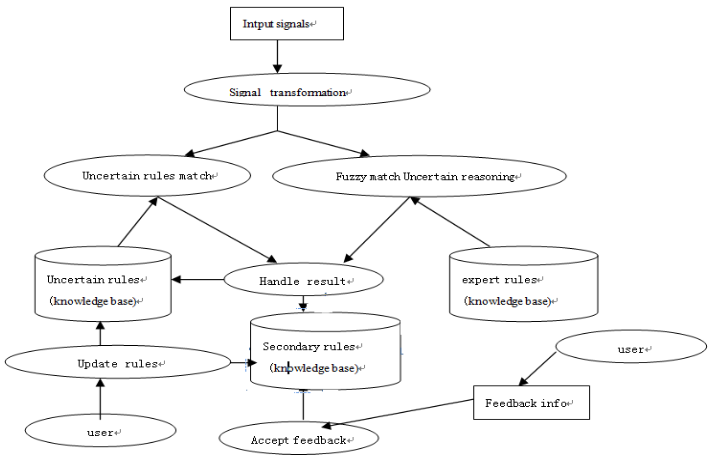


Figure 3. Uncertain Decision Making reasoning flow diagram

3.3 Decision Tree

The decision tree is a data mining method to find out classification of data set by constructing decision tree. The key part of it is how to construct highly-accurate size of a small tree. As a tree structure similar to flow chart, the decision tree has the best extension Property as its internal node and category Property value as its leaf node. The edge of internal node is the value of best extension Property. The data sets of the internal nodes are impure and classified to different categories. Each internal node means the diagnosis of one Property. The data set of the root node is training set. The data sets of other internal nodes are the training subsets. The data set of the leaf node is pure training subset, which represents one diagnosis output. When decision tree is being constructed, many branches, which indicate the noise or outlier in the training data, will be tested and cancelled by Tree pruning method (J.Ross et al., 1997) to better the accuracy of the unknown data classification.

In this system, the decision tree reasoning module has the capability to construct reasoning functional tree structure through massive training samples. Every internal node represents the diagnosis on one property and the branch of such a node means each result of the diagnosis, as a result, every leaf node means one diagnosis consequence. In order to classify and indentify the unknown signal data, the property value in the data sets of the decision tree is to be diagnosed from the root node to the leaf node, which forms a route of the category reasoning on the relative object. If the conclusion cannot be reached one at a time according to the expert rules, the already-constructed decision can be used to do reasoning to bring out a somewhat reliable conclusion. In addition, the decision reasoning module can give feedback to the deduced rules, save the feedback and perform retraining to continuously improve the accuracy of the reasoning mechanism. The following, as shown in figure 6, is the flow diagram.

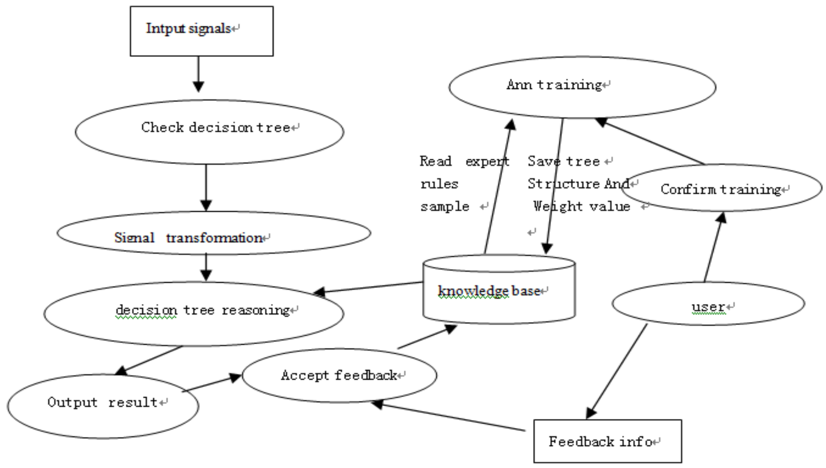


Figure 4. Decision Tree reasoning flow diagram

3.4 Research and Evaluation of Multi-Reasoning Mechanism

Reasoning Mechanism needs to be evaluated by a specific method. Reasoning mechanism model aims to correctly classify a new sample and the measure of mechanism quality is common reasoning accuracy. All available samples are divided into two types: training samples and test samples. First, the training samples are used to construct reasoning model, and then the test samples are used to evaluate the model on the basis of the accuracy. Training samples and test samples should be not only large enough but also independent. With the training samples and test samples, it is necessary to decide which statistical method to measure the accuracy of the forecasts. In this paper, we use Matthew's Correlation Coefficient (Smith, 1998), as shown in equation 1: TP_i denotes the number of correct classification, FP_i denotes Classification faults, FN_i denotes Negative faults, TN_i denotes The number of correct negative.

$$C_i = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FN)(TP + FP)(TN + FP)(TN + FN)}} \quad (1)$$

Matthew correlation coefficient is between number 1 and -1. If the value is 1, then the model is completely positive correlation, -1 indicates perfect negative correlation predicted, 0 means no correlation. If the reasoning mechanism of Matthew correlation coefficient is as close to 1 as possible, the accuracy of this reasoning will be better. For instance, as a reasoning consequence 'replacement accelerometer' of clause 10, the historical data is collected from 100 diagnostic records to calculate the Matthew correlation coefficient, the results shown in Table 2.

Table 2. Matthew's Correlation Coefficient of clause 10 "Change Accelerometer"

	Ann	Uncertain decision	Decision tree
TP	30	28	29
FP	1	2	2
FN	0	2	1
TN	68	68	68
C1	0.954	0.936	0.941

Similarly, wit h th e rem aining reasoning consequence of clause 10, we calc ulated th e Matth ew correlation coefficient, as shown in Table 3.

Table 3. Matthew's Correlation Coefficient of diagnosis clause 10

(clause 10) diagnosis consequence	Ann	Uncertain decision	Decision tree
Change Accelerometer	0.954	0.936	0.941
Return to factory	0.956	0.932	0.938
replace 1A4A1 board	0.967	0.927	0.944
replace R34	1.000	0.942	0.945
replace 1A4A10	0.974	0.933	0.947
reset 1A4A8	0.966	0.926	0.951
$\overline{C}$	0.970	0.933	0.944

From the analysis of data showed 10 diagnosis clause, three reasoning mechanisms are different in accuracy, and artificial n eural n etwork is better. Dia gnosis clau se 11 and 14 ar e in troduced to do co mparison study to gether with diagnosis clause 10, as shown in Table 4, to build mechanism compatible with this system.

Table 4. Description of diagnosis clause 10,11 and 14

diagnosis clauses	Ann	Uncertain decision	Decision tree
clause 10	0.970	0.933	0.944
clause 11	0.974	0.952	0.986
clause 14	0.975	0.992	0.947

The a verage Matthew correlation coe fficients of diagnosis clause 11 and 14 a re respectively c alculated in accordance with the method of diagnosis clause 10, as shown in table 5.

Table 5. Comparison of average Matthew's Correlation Coefficient

diagnosis clauses	Signal number	Rule number	Consequence number
clause 10	6	6	6
clause 11	11	14	13
clause 14	18	33	28

We can see from Table 5, none of three reasoning mechanisms has the highest reasoning accuracy on all diagnosis clauses, which means that concerning the issues studied in this system, no single mechanism is superior to other two mechanisms. In this system we uses a common technology to combine the artificial neural network, uncertain reasoning and decision tree, so as to create an improved mechanism to increase the reasoning accuracy.

In this system the growing voting Mechanism is used to set the credit weight value of three reasoning mechanisms for each clause and constantly adjust the weight values to make the most successful reasoning mechanism in history get higher priority.

Take clause 14 for example, first set the reasoning mechanism for each initial value of 1, and then adjust the weight values according to results of each reasoning process through feedback and validation. The current value of uncertainty decision is 152, the current value of ANN is 138, the current value of decision tree is 115, when a new group of signals is being reasoned, the result of the uncertainty rule is 'replace 4A1A board', the result of artificial neural network is 'replace 4A4B board', the result of the decision tree is 'reset 1A4A8'. In this case, the voting weights will be calculated:

The value of 'replace 4A1A board' $= 152/(152+115+138) = 37.53\%$

The value of 'replace 4A1B board' $= 138/(152+115+138) = 34.07\%$

The value of 'reset 1A4A8' $= 115/(152+115+138) = 28.39\%$

When user gives feedback on this diagnosis reasoning, if the actual failure is 'replace 4A1A', the value of uncertainty decision will be added by 2; if the actual failure is 'replace 4A1B', the value of artificial neural network will be added by 2; If the actual failure is 'reset 1A4A8', the value of decision tree will be added by 2; if the actual failure turns out to be 'replace 3A1A', not among the above three consequences, the value of three reasoning mechanisms will be reduced by 1. The current credit weight values of the diagnosis clause 10, 11 and 14 are shown in Table 6:

Table 6. Credit weight values of different reasoning mechanism

MachType	Diagnosis Clauses	reasoning mechanism	Credit weight value	LastUpdate
1	10	Uncertain decision	198	2005-8-6
1	10	ANN	212	2005-8-6
1	10	Decision tree	186	2005-8-6
1	11	Uncertain decision	124	2005-9-12
1	11	ANN	152	2005-9-12
1	11	Decision tree	174	2005-9-12
1	14	Uncertain decision	152	2005-9-12
1	14	ANN	138	2005-9-12
1	14	Decision tree	115	2005-9-12

After a period of training, the reasoning mechanism with the higher weight value (that is, the higher success rate) will become more important and reliable, and taking the contribution to the reasoning process of other mechanisms into full consideration, multi-Reasoning mechanism can come to a more accurate final reasoning result. The flow diagram is shown in figure 6.

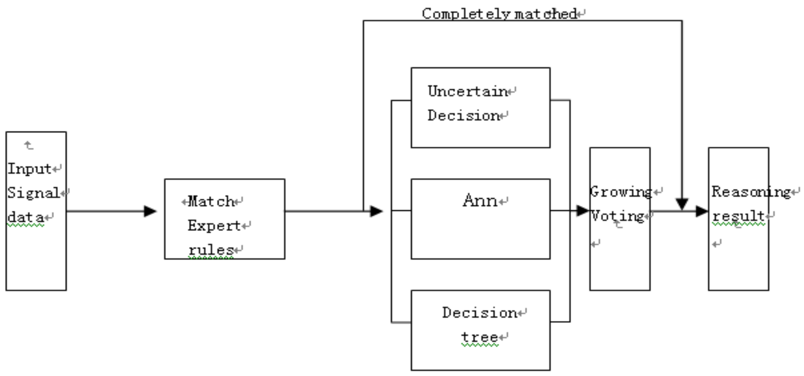


Figure 5. Multi-Reasoning Mechanism flow diagram

The average Matthew correlation coefficients, calculated through the same historical data in table 5, are shown in Table 7. As a result, we can see the growing voting mechanism, which combines with different mechanisms, improves the accuracy of reasoning in port machine diagnosis by data analysis.

Table 7. Comparison of average Matthew's Correlation Coefficient including multi-Reasoning mechanism

Diagnosis Clauses	Ann	Uncertain decision	Decision tree	Growing voting
Clause 10	0.970	0.933	0.944	0.982
Clause 11	0.974	0.952	0.986	0.986
Clause 14	0.975	0.992	0.947	0.994

4. CONCLUSION

Based on the theories above, the port machine diagnosis expert system has been developed. The system consists of four parts, as shown in Figure 7, signal acquisition server, knowledge base management server, port machine diagnosis platform and knowledge database. System process flow is as follows: diagnosis platform receives the signals from signal acquisition server, with connecting to knowledge base of specific type of port machine, multi- reasoning mechanism starts work to achieve reliable diagnosis results. The knowledge base management server is used to maintain knowledge and train reasoning mechanism to improve the accuracy of this system.

This paper is just one feasible proposal for the complexity port machine diagnosis. It needs to track the development of cutting edge to further optimize the reasoning mechanism, so as to improve the accuracy and performance of the system to achieve better diagnosis results.

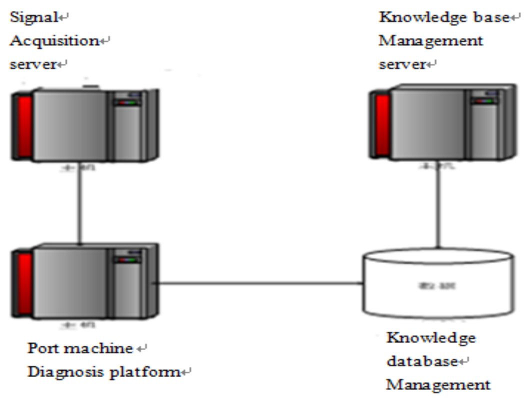


Figure 7. system structure

5. REFERENCES

S.Horikawa.On fuzzy modeling using fuzzy neural network with BP algorithm. IEEE Trans.Neural Networks，2002，2：30-34.

Timo Sorsa, Heikkin.Koivo. Application of Artificial Neural Networks in Process Fault. Diagnosis.Automatica，2003 ，29：834-848.

J.Ross Quinlan. C4.5： Programs For Machine Learning. IEEE Trans，1997， 3：31-34.

Mark Brodie, Irina Rish. Intelligent Probing： a Cost-Effective Approach to Fault Diagnosis in Computer Networks. IBM Systems Journal，2002，41：103-107.

E.Barnard.Optimization for training neural nets. IEEE Transations on Neural Networks，1992，3(2)：232-240.

Shyi, Zhong Zhi, Zhou Han. Applying Case-based Reasoning to Engine Oil Design. Artificial Intelligence in Engineering，1996，12：167-172.

Smith D.J .A Artificial Intelligence-Today’s New Design and Diagnostic Tool. Power Engineering，1998，93(1)：26-30.

Stuart Russell. Artificial Intelligence—A Modern Approach(Second Edition). 2004：151-288，355-412.

Joseph Giarratano. Expert System principles and programming. 2002：1-39，40-158.

Session B3: Seaport and Transportation 2

·Day1: Sep. 15 (Wed.)

·Time: 13:20 - 14:40

·Chair: Chulung Lee

·Room: Iris, 4F

REPOSITIONING EMPTY CONTAINER REPOSITIONING IN SURPLUS PORT AREA

Young Jin Han[1], Quang Vinh Dang[2], Won Young Yun[3*]
[1]Department of Industrial Engineering
Pusan National University,
30, Jangjeon-Dong, Geumjeong-Gu,
Busan, South Korea
E-mail: shoot78@pusan.ac.kr[1], vinhise@pusan.ac.kr[2], wonyun@pusan.ac.kr[3*]

1. MODEL DESCRIPTION

Empty containers are important logistics resources need to transport freight from ports to ports or in inland transportation. To improve customer-service levels and productivity, shipping companies wish to manage and operate empty containers efficiently. Owing to the trade imbalances, some ports have a large number of unnecessary empty containers but other ports need more empty containers. There exists the same problem among depots in inland transport systems due to demand and supply imbalance of empty containers. In order to solve the imbalance problem, there are three options which are repositioning, ordering of empty containers and leasing of empty containers to replenish empty containers. But ordering of empty containers requires a long lead-time, leasing of empty containers should be also returned to leasing companies after a specified period.

On the other hand, shipping companies desire to move empty containers in surplus areas to shortage areas to reduce shortage and holding costs. However repositioning of empty containers between depots and ports needs a lead-time and transportation cost. Therefore, we need efficient policy for repositioning empty containers to minimize total cost.

Optimization studies about empty container management have been done by several researchers. Cheung and Chen (1998) considered a two-stage stochastic network for the dynamic empty containers allocation problem. They used a single-commodity stochastic and dynamic network to model both owned empty containers and leased containers. Thus, the two-stage stochastic network has been solved using the stochastic linearization method. . Li et al. (2004) developed a (*u*, *d*) policy to deal with the allocation of empty containers and built a mathematical model to compare the average costs of finite horizon with infinite horizon case. This study was then extended by Li et al. (2007) in case of multi-ports. They concerned with the allocation of empty containers between multi ports and developed a heuristic algorithm to minimize the average cost.

In this study, we consider an inventory control problem of empty containers in two ports with two depots in surplus area. There are depots, ports and customers in the inland transport network shown in Figure 1. Shipping company operates two-depots and ports to store empty containers and provide them for transportation of freight. For transporting freight, depots and ports send empty containers to their customers. Also depots and ports usually send the remaining empty containers to shortage areas and reduce holding and shortage costs. In general, the demands of empty containers in customer places may be uncertain and it is very difficult to know the exact amount of necessary empty containers. We assume that demands of empty containers in customer locations per unit time are independent and identically distributed random variables. We want to determine the amount of repositioning empty containers and the safety stock level. The holding, transportation and shortage costs are considered. We propose an efficiency policy to determine the reasonable repositioning amount of empty containers for minimizing total expected cost.

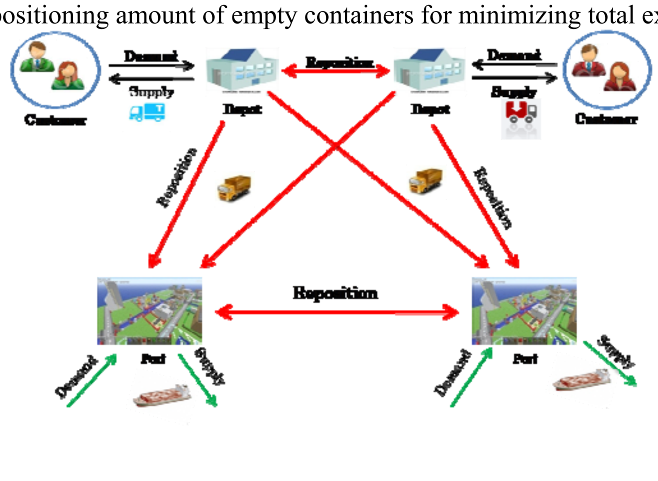


Figure 1. Inland transportation system for empty containers

2. EMPTY CONTAINER REPLENISHMENT POLICY

2.1 Assumptions

In this paper, an inventory control problem for empty containers is studied under the following assumptions:

- There are two depots and ports.
- Single commodity, 40ft containers, is considered.
- Holding, transportation and shortage costs are given.
- Net dem ands of em pty cont ainers per unit t ime peri od a re i ndependent an d i dentically di stributed random variables.
- The lead times for inland positioning between depots and ports are different.

2.2 Notations

The notation used in the inventory model is as follows:

t : Discrete time-period.
i : Index of depots
j : Index of ports
I_D^i : Inventory position at depot *i*.
I_P^j : Inventory position at port *j*.
PO_D^i : Number of empty containers that can be positioned out from depot *i*.
PI_D^i : Number of empty containers that can be positioned in at depot *i*.
PI_P^j : Number of empty containers that can be positioned in at port *j*.
A_D^t : Set of depots that empty containers can be positioned in.
A_P^t : Set of ports that empty containers can be positioned in.
B_D^t : Set of depots that empty containers can be positioned out.
B_P^t : Set of ports that empty containers can be positioned out.
C_D^t : Set of depots that empty containers cannot bet either positioned out or in.
C_P^t : Set of ports that empty containers cannot be either positioned out or in.

The decision variables:

S_D^i : Safety stock of depot *i*.

S_P^j : Safety stock of port *j*.

2.3 Empty containers replenishment policy

Based on the estimated inventory position of depots and ports we determine safety stock level and repositioning amount of empty containers. At each time period *t*, we estimate the inventory position $I_D^i(t)$ of depots and $I_P^j(t)$ of ports.

Step 1. Classify depots and ports into three sets.

- First set includes depots and ports which can reposition out empty containers.

$$A_D^t = \{I_D^i(t) > S_D^i\}, \quad A_P^t = \{I_P^j(t) > S_P^j\}$$

- Second set includes depots and ports which can reposition in empty containers

$$B_D^t = \{I_D^i(t) < S_D^i\}, \quad B_P^t = \{I_P^j(t) < S_P^j\}$$

- Final set includes depots and ports which not either reposition out or in empty containers.

$$C_D^t = \{I_D^i(t) < S_D^i\}, \quad C_P^t = \{I_P^j(t) < S_P^j\}$$

Step 2. Reposition empty containers between depots

Step 2.1 Empty containers are positioned one by one from an origin depot in set A where has the largest number of empty containers positioned out to a destination depot in set B where has the largest number of empty containers positioned in.

Step 2.1 The numbers of empty containers that can be positioned out and in are updated.

$$\sum_{i=1}^{n} PO_D^i(t) > 0, \; \sum_{i=1}^{n} PI_D^i(t) > 0 \quad \text{Go to Step 2.1}$$

$$\sum_{i=1}^{n} PO_D^i(t) > 0, \; \sum_{i=1}^{n} PI_D^i(t) = 0 \quad \text{Go to Step 3}$$

The repositioning of empty containers between depots continues until either the total of $PO_D^i(t) = 0 \;\; (i \in A)$ or $PI_D^i(t) = 0 \;\; (i \in B)$ equals zero.

Step 3. Reposition empty containers from depots to ports

Step 3.1 Empty containers are repositioned one by one from an origin depot in set A to a destination port in set B where has the largest numbers of empty containers that can be positioned in.

Step 3.2 The numbers of empty containers that can be positioned out and in are updated.

$$\sum_{i=1}^{n} PO_D^i(t) > 0, \; \sum_{i=1}^{n} PI_P^j(t) > 0 \quad \text{Go to Step 3.1}$$

$$\sum_{i=1}^{n} PO_D^i(t) > 0, \; \sum_{i=1}^{n} PI_P^j(t) = 0 \quad \text{Go to Step 1 after } t+1$$

The repositioning of empty containers between depots continues until either the total of $PO_D^i(t) = 0 \;\; (i \in A)$ equals zero.

ACKNOWLEDGMENT

"This work was supported by the Grant of the Korean Ministry of Education, Science and Technology" (The Regional Core Research Program/Institute of Logistics Information Technology).

3. REFERENCES

Dang, Q.V. (2010). Replenishment Policies for Empty Containers in an Inland Multi-depot System, Master Thesis, Pusan National University, South Korea.

Chenung, R.K. and Chen, C.Y. (1998). A Two-stage Stochastic Network Model and Solution Methods for the Dynamic Empty Container Allocation Problem, *Transportation Science, 32:* 142-166.

Li, J.A., Liu, K., Leung, S.C.H. and Lai, K.K. (2004). Empty Containers Management in a Port with Long-run Average Criterion, *Mathematical and Computer Modeling, 40:* 85-100.

Urban, T.L. (2005). A Periodic-review Model with Serially Correlated, Inventory-level-dependent Demand, *International Journal of Production Economics, 95:* 287-295.

Li, J.A., Wu, Y. and Liu, K. (2007). Allocation of Empty Containers between Multi-depots, *European Journal of Operation Research, 182:* 400-412.

Law, A.M. and Kelton, W.D. (2000). Simulation Modeling and Analysis, 3rd ed., McGraw-Hill, New York, USA.

Silver, E.A., Pyke, D.F. and Peterson, R. (1998). Inventory Management and Production Planning and Scheduling, 3rd ed., John Wiley & Sons, New York, USA.

OPTIMAL PRICING AND GUARANTEED LEAD TIME WITH CONSIDERATION OF LATENESS PENALTIES

Ki-sung Hong [1], Chulung Lee [2*]

[1] Division of Information Management Engineering
Korea University
Anamdong 5-ga, Seongbuk-gu, 136-713
Seoul, Republic of Korea
e-mail: justlikewind@korea.ac.kr

[2] (Corresponding Author) Division of Information Management Engineering
Korea University
Anamdong 5-ga, Seongbuk-gu, 136-713
Seoul, Republic of Korea
e-mail: leecu@korea.ac.kr

Abstract: This paper studies a price and guaranteed lead time decision for a supplier that offers a fixed guaranteed lead time for the product. If the supplier is not able to meet the guaranteed lead time, the supplier must pay the lateness penalty to the customers. Thus, the expected demand is a function of price, guaranteed lead time and lateness penalty. We first develop a mathematical model for a given supply capacity to determine the optimal price, guaranteed lead time and lateness penalty to maximize the total profit. We then consider the case where it is also possible for the supplier to increase a capacity marginally and compute the optimal capacity expansion.

1. INTRODUCTION

The increased competition has affected service providers and manufacturers to introduce new products in market and time has evolved as the competitive paradigm (Blackburn 1991, Hum and Sim 1996). As time has become a key for business success, lead time reduction has emerged as a key competitive edge in service and manufacturing (Van Beek and Van Putten 1987, Suri 1998, Hopp and Spearman 2000). This new competitive paradigm is termed as time-based competition.

Suppliers exploit customers' sensitivity to time to increase prices in return for a shorter lead time. For instance, Amazon.com charges more than double the shipping costs to guarantee delivery in two days, while its normal delivery time may be as long as a week (Ray and Jewkes, 2004). Likewise suppliers differentiate their products based on lead times in order to maximize the supplier's revenue (Boyaci and Ray, 2003). On this case, the lead time reduction has provided suppliers with new opportunities. Additionally, in today's world of global economy, suppliers are increasingly depending on fast response times as an important source of sustainable competitive advantage. As a result, to consider the influence of lead times on demand is need.

This paper studies a supplier that is using the guaranteed lead time to attract customers and supply a product in a price and time sensitive market. The supplier must pay the lateness penalty to the customers whenever the actual lead time exceeds the guaranteed lead time. The price, guaranteed lead time and lateness penalty are decision variables of our model.

The time-based competition is first studied by Stalk and Hout (1990) who address the effect of time as a strategic competitiveness. So and Song (1998), Palaka et al. (1998), Ray and Jewkes (2004) and Hill and Khosla (1992) have also studied the optimal pricing and delivery time decisions, while modeling the supplier's operations as a single server queue. So and Song (1998) study the impact of using delivery time guarantees as a competitive strategy in service industries where demands are sensitive to both price and delivery time. They use an M/M/1 queueing model and propose a mathematical framework to understand the interrelations among the pricing, delivery time.

Palaka et al. (1998) use a similar framework but a linear relationship between the mean demand, price and the delivery time and examine the lead time setting, capacity utilization, and pricing decision. Ray and Jewkes (2004) further extend this study by modeling price as a function of delivery time, besides demand being a function of price and delivery time. Hill and Khosla (1992) also study a similar tradeoff between price and delivery time but in a deterministic framework. So (2000), Tsay and Agarwal (2000), Allon and Federgruen (2008) and Pekgun et al. (2006) also study similar problems but their models do not study the lateness penalty decision.

The primary objective of this research is to develop a model to determine the price, the guaranteed lead time and the lateness penalty for the product to maximize the total profit. We first develop a model for a given capacity and then, we consider the case where it is also possible for the supplier to increase the capacity marginally and develop a mathematical model to determine the optimal capacity expansion.

The rest of this paper is organized as follows: In Section 2, we formulate the initial model to determine the price, the guaranteed lead time and the lateness penalty for the objective of maximizing the total profit. Optimal properties for the model are obtained and an efficient algorithm is provided to compute optimal price, guaranteed lead time and lateness penalty. We then extend our results in Section 3 to consider the case where to increase the capacity is possible. To gain further insight, we conduct computational experiments and analyze the effect of service level on optimal decision and the sensitivity of the optimal solution with respect to the model parameters in Section 4. Our conclusions are provided in Section 5.

2. THE MODEL FOR THE FIXED CAPACITY CASE

2.1 Assumptions and notations

We consider a supplier that serves customers in a price and time sensitive market and assume that the supply capacity, μ, is fixed. The supplier offers a fixed guaranteed lead time and a service reliability level to the customer. The price for the product, guaranteed lead time and service reliability level, are denoted by p, l and s ($0 \le s \le 1$), respectively. The demand is modeled as a Poisson process, and the service time of the customer orders are assumed to be exponentially distributed. The customers are served on a first-come first-served basis. These assumptions allow us to use an M/M/1 queueing model of the supplier's operations. In this case, the supplier needs to optimally decide on the price, guaranteed lead time and service reliability level.

The supplier's cost structure includes the following two main categories; direct unit variable costs and lateness penalty costs. The direct unit variable costs mean all costs that are proportional to production volume, such as the cost of direct materials and labor. The lateness penalty costs mean direct compensation paid to customers for not meeting the guaranteed lead time, and we assumed that the lateness penalty is constant regardless of how late the product arrives (e.g. Domino Pizza's 30 minute or free delivery policy). The direct unit variable cost and lateness penalty cost for a product are denoted by m and π, respectively.

The form of our demand model is similar to Boyaci and Ray (2006). The demand model is a standard linear model where mean demand is decreasing in the price and the lead time, and increasing in the service reliability level and the lateness penalty. That is,

$$\lambda(p,l,s,\pi) = a - bp - cl + ds + e(1-s)\pi \qquad (1)$$

where:

λ : expected demand for the product

a : demand corresponding to zero price, zero guaranteed lead time, zero service reliability level and zero lateness penalty

b : price sensitivity of demand

c : lead time sensitivity of demand

d : service reliability level sensitivity of demand

e : lateness penalty sensitivity of demand

2.2 Mathematical model

We assume that the supplier's objective is to maximize the expected total annual profit which can be expressed as:

$$\prod(p,l,s,\pi) \equiv (p-m)\lambda(p,l,s,\pi) - \frac{\pi\lambda(p,l,s,\pi)}{\mu - \lambda(p,l,s,\pi)} e^{-(\mu-\lambda)l} \qquad (2)$$

In the objective function, $(p-m)\lambda$ represents the expected annual revenue for the product. Since we assume that an M/M/1 queueing system and the lateness penalty is constant regardless of the length of lateness, the expected annual lateness penalty for the product is expressed as (lateness penalty) $\times$ (expected number of customers in system) $\times$ (probability that a actual lead time exceeds the guaranteed lead time) and given by $\frac{\pi\lambda}{\mu-\lambda}e^{-(\mu-\lambda)l}$. Thus, the mathematical model is expressed as:

$$\underset{p,l,s,\pi}{Maximize} \quad (p-m)\lambda(p,l,s,\pi)-\frac{\pi\lambda(p,l,s,\pi)}{\mu-\lambda(p,l,s,\pi)}e^{-(\mu-\lambda)l} \tag{3}$$

$$Subject\ to \quad \lambda \le \mu \tag{4}$$

$$1-e^{-(\mu-\lambda)l} > s \tag{5}$$

$$0 \le p,\ 0 \le l,\ 0 \le \pi \tag{6}$$

In this mathematical model, constraint (4) is system stability constraints that the supplier's mean service rate exceeds the mean demand rate. Constraint (5) represent that actual service level is larger than service reliability level. Constraint (6) restricts the price, guaranteed lead time and lateness penalty for the product to non-negative values.

Observation 1. The service level constraint (4) is non-binding if and only if the service reliability level, s, is strictly less than the critical value, $s_c = 1 - c/b\pi$, and otherwise, the service level constraint (4) is binding.

Proof. From Equation (1), we obtain $p = \frac{1}{b}\left[a - cl + ds + e(1-s)\pi - \lambda\right]$.

Substitute $p = \frac{1}{b}\left[a - cl + ds + e(1-s)\pi - \lambda\right]$ in Equation (3). Thus, the mathematical model is expressed as:

$$\underset{\lambda,l,\pi}{Maximize} \quad \left[a-cl+ds+e(1-s)\pi-\lambda-bm\right]\frac{\lambda}{b}-\frac{\pi\lambda}{\mu-\lambda}e^{-(\mu-\lambda)l} \tag{7}$$

Applying the Lagrangian multiplier method, a stationary point to the problem must satisfy:

$$\frac{d\prod(\lambda,l,\pi)}{d\lambda}=0\ ,\quad \frac{d\prod(\lambda,l,\pi)}{dl}=0\ ,\quad \frac{d\prod(\lambda,l,\pi)}{dR}=0\ ,\quad \frac{d\prod(\lambda,l,\pi)}{d\gamma}=0\ ,\quad \gamma>0$$

where

$$\prod=\left[a-cl+ds+e(1-s)\pi-\lambda-bm\right]\frac{\lambda}{b}-\frac{\pi\lambda}{\mu-\lambda}e^{-(\mu-\lambda)l}+\gamma\left[(\mu-\lambda)l-\log\left(\frac{1}{1-s}\right)\right]$$

Therefore, it must be true for any optimal solution that:

$$\begin{aligned}\frac{d\prod(\lambda,l,\pi)}{dl}=0 &\Leftrightarrow 1-\frac{c}{b\pi}+\frac{\gamma(\mu-\lambda)}{\pi\lambda}=1-e^{-(\mu-\lambda)l}\\ &\Leftrightarrow s_c+\frac{\gamma(\mu-\lambda)}{\pi\lambda}=1-e^{-(\mu-\lambda)l}\end{aligned} \tag{8}$$

We first demonstrate that if the service level constraint is non-binding, then $s_c > s$ (where $s_c = 1-\frac{c}{b\pi}$). Note that if the service level constraint is non-binding, then $\gamma = 0$. Since $\gamma = 0$, Equation (8) implies that $s_c = 1-e^{-(\mu-\lambda)l}$. Also, if the service level constraint is non-binding, then $s_c = 1-e^{-(\mu-\lambda)l} > s$ and service level is s_c. We now demonstrate that if the service level constraint is binding then $s_c \le s$ and the service level is s. If the service level constraint is binding then $1-e^{-(\mu-\lambda)l}=s$. Combining this with Equation (8) above implies that $s_c+\frac{\gamma(\mu-\lambda)}{\pi\lambda}=s$. Since γ, $(\mu-\lambda)$, π, $\lambda\geq 0$, we conclude that $s_c\leq s$ which proves our result. □

Proposition 1. The upper bound of π is equivalent to $\overline{\pi}=c/[b(1-s)]$ and if the service level constraint is non-binding, then the optimal lateness penalty, π^* is given by $\pi^*=c/[b(1-s)]$.

Proof. From Observation 1, the probability that the supplier is not able to meet the guaranteed lead time $e^{-(\mu-\lambda)l}$, is equivalent to $1-Max[s_c,s]$. If $s\geq s_c$, then $e^{-(\mu-\lambda)l}=1-s$ and $\pi\leq c/[b(1-s)]$, and if $s<s_c$, then $e^{-(\mu-\lambda)l}=1-s_c$, $\pi>c/b(1-s)$ and $\lambda=a-bp-cl+ds+ce/b$ so that π does not affect the demand. In this case, the objective function (2) is a non-increasing function of π so that the optimal lateness penalty, π^* is equivalent to $c/[b(1-s)]$. Thus, we conclude that the upper bound of π is equivalent to $\overline{\pi}=c/[b(1-s)]$. □

Proposition 2. For a given values of p and l, the objective function (3) is a concave function of π.

Proof. Taking the second order derivative of (3) with respect to π, we have

$$-\frac{[2d(1-s)(\mu-\lambda)+d(1-s)(a-bp-cl)]+2d(1-s)[(a-bp-cl)+2d(1-s)\pi(\mu-\lambda)+\theta\pi\lambda_1]}{(\mu-\lambda)^3}e^{-(\mu-\lambda)l}$$
$$-\frac{2d(1-s)l[(a-bp-cl)+2d(1-s)\pi(\mu-\lambda)+d(1-s)\pi\lambda]}{(\mu-\lambda)^2}-\frac{d(1-s)l\pi\lambda}{\mu-\lambda}e^{-(\mu-\lambda)l}<0 \quad (9)$$

Since the second order derivative is always less than zero, the objective function (3) is concave with respect to π. □

Proposition 3. For a given value of π,

1. The optimal guaranteed lead time, l^* is given by $\log k/(\mu-\lambda^*)$ in both the binding and non-binding cases.
2. The optimal price, p^*, is given by $[a-\lambda^*-cl^*+ds-e\pi/x]/b$.

Proof. From Proposition 1, we know that the service level will be given by $Max\{s,s_c\}$. Therefore,

$$1-e^{-(\mu-\lambda)l^*}=Max\{s_c,s\}$$
$$\Leftrightarrow l^*=\frac{1}{(\mu-\lambda)}\log\left(\frac{1}{1-Max\{s_c,s\}}\right)\Leftrightarrow l^*=\frac{1}{(\mu-\lambda)}\log x \quad (10)$$

where $x=Max\{1/(1-s),b\pi/c\}$.

Substitute the above expression and $e^{-(\mu-\lambda)l^*}=1/x$ into the Equation (6), we obtain follow equation:

$$\prod(\lambda)=\left[a-\frac{c\log x}{(\mu-\lambda)}+\frac{d\pi}{x}-\lambda-bm\right]\frac{\lambda}{b}-\frac{1}{x}\frac{\pi\lambda}{\mu-\lambda} \quad (11)$$

Differentiating $\prod(\lambda)$ with respect to λ, we obtain the first and second order condition:

$$\prod{}'(\lambda)=\left[\frac{a}{b}+\frac{d\pi}{bx}-m-\frac{2\lambda}{b}\right]-\frac{\mu}{(\mu-\lambda)^2}\left[\frac{c\log x}{b}+\frac{\pi}{x}\right] \quad (12)$$

$$\prod{}''(\lambda) = -\frac{2}{b} - \frac{2\mu}{(\mu-\lambda)^3}\left(\frac{c\log x}{b} + \frac{\pi}{x}\right) < 0 \tag{13}$$

Since second order condition is always less than zero, $\prod(\lambda)$ is concave with respect to λ. Thus, we calculate the optimal solution, λ^* to maximize the $\prod(\lambda)$, and optimal guaranteed lead time, l^* is given by $\log x/(\mu-\lambda^*)$ in both binding and non-binding cases.

To prove part (b), we note that $\lambda^* = a - bp^* - cl^* + ds - e\pi/x$ and solve for p^*. □

Using the above optimality conditions, we can develop an iterative algorithm to compute the optimal solution. In general, iterative algorithms converge to local optimal solutions. To avoid this problem, Kaspi and Rosenblatt (1991) developed the RAND algorithm. They obtained several local optimal solutions from the iterative algorithm with different initial values, and selected the best solution among all the local optimal solutions. Their simulation studies show that the initial value has a strong influence on the quality of the obtained solution. Using this idea, we modify the RAND algorithm and develop a new algorithm. The modified RAND algorithm for the fixed capacity case, denoted as SRA-RAND_FC, proceeds as follows.

Modified RAND algorithm for fixed capacity case (SRA-RAND_FC)

Step 1. For given *n*, divide [0, $\frac{c}{b(1-s)}$] into *n* equally spaced values of π_j : (π_1, ⋯, π_j, ⋯, π_n). Set *j* = 0.

Step 2. Set *j* = *j*+1 and *r* = 0. Let $\pi_j(r) = \pi_j$.

Step 3. Set *r* = *r*+1.

Step 4. For a given value of $\pi_j(r-1)$, compute p_j and l_j.

Set $p_j(r) = p_j$ and $l_j(r) = l_j$.

Step 5. For given values of $p_j(r)$ and $l_j(r)$, compute π_j.

Set $\pi_j(r) = \pi_j$.

Step 6. If $\pi_j(r) \neq \pi_j(r-1)$, go to Step 3.

Otherwise, set $p_j^*(r) = p_j(r)$, $l_j^*(r) = l_j(r)$ and $\pi_j^*(r) = \pi_j(r)$.

Compute $\prod_j$ for this $[p_j^*(r), l_j^*(r), \pi_j^*(r)]$.

Step 8. If *j* ≠ *n*, go to Step 2.

Otherwise, stop and select $[p_j^*(r), l_j^*(r), \pi_j^*(r)]$ with the maximum $\prod_j$.

We will now provide an illustration of the SRA-RAND_FC process using the numerical example, as that shown in Table 1. In Step 1, we obtain $\bar{\pi}$ = 4.0. When n =5, the solutions for the various iterations of SRA-RAND_FC are shown in Table 2.

Table 1. The parameter values for numerical example

parameter	value	parameter	value
a	50	μ	10
b	2	s	0.75
c	2	$\bar{\pi}$	4.0
d	1	n	5
m	3		

Table 2. Summary of the SRA-RAND_FC solution procedure

π_j	Iteration *r*	p_j	l_j	new π_j	TP_j
π_1=0.8	1	9.96027	0.529504	0.8008	34.9245
	2	9.96027	0.529504	0.8008	**34.9245**
π_2=1.6	1	10.1672	0.491299	1.5898	34.0584
	2	10.1672	0.491313	1.5913	34.0581
	3	10.1673	0.491294	1.5949	34.0577
	3	10.1673	0.491294	1.5949	34.0577
π_3=2.4	1	10.3584	0.461614	2.40028	33.2977
	2	10.3584	0.461614	2.40028	33.2977
π_4=3.2	1	10.5327	0.438458	3.1996	32.6439
	2	10.5327	0.438458	3.1996	32.6439
π_5=4.0	1	10.6967	0.41955	3.9972	32.0649
	2	10.6967	0.41955	3.9972	32.0649
p^*=9.96027, l^*=0.529504, π^*=0.8008					

3. THE MODEL FOR THE CAPACITY EXPANSION CASE

We now consider the case where it is also possible for the supplier to increase the capacity marginally. To consider the case, the following additional notations are employed.

z : fractional increase in the processing rate
$\bar{z}$: upper bound on capacity expansion
k : cost of increasing the processing rate by one unit capacity

The type of capacity expansion that we consider is one of a relatively short term nature, e.g., hiring part-time or temporary workers, running overtime (Palaka et al., 1998). With these types of capacity expansion, it is reasonable to assume that we can model the capacity expansion as a continuous decision variable and use a linear cost of additional capacity. Thus, we formulate our capacity expansion model as:

$$\underset{p,l,s,\pi,z}{Maximize} \quad (p-m)\lambda - \frac{\pi\lambda}{\mu(1+z)-\lambda} e^{-[\mu(1+z)-\lambda]l} - k\mu z \tag{14}$$

$$Subject\,to \quad \lambda \le \mu(1+z) \tag{15}$$

$$1-e^{-[\mu(1+z)-\lambda]l} > s \tag{16}$$

$$0 \le z \le \bar{z} \tag{17}$$

$$0 \le p,\ 0 \le l,\ 0 \le \pi \tag{18}$$

Note that Proposition 1 and 2 are valid for the above model. The following proposition characterizes the optimal decisions for the capacity expansion model

Proposition 4. For a given value of π,

(1) The optimal capacity expansion, z^* is given by $1/\mu(\lambda^* + \sqrt{\lambda^* G/bk}) - 1$ where $G = c\log x + b\pi / x$.

(2) The optimal guaranteed lead time, l^* is given by $\log x / (\mu(1+z^*) - \lambda^*)$

(3) The optimal price, p^*, is given by $[a - \lambda^* - cl^* + ds - e\pi / x]/b$.

Proof. The proof of Proposition 5 is similar to that of Proposition 4. We begin by developing the expression for the optimal lead time. From proposition 1, the service level will be given by $Max[s_c, s]$. Therefore,

$$1 - e^{-(\mu(1+z)-\lambda)l^*} = Max\{s_c, s\}$$

$$\Leftrightarrow l^* = \frac{1}{(\mu(1+z)-\lambda)}\log\left(\frac{1}{1-Max\{s_c,s\}}\right) \Leftrightarrow l^* = \frac{\log x}{(\mu(1+z)-\lambda)} \tag{19}$$

To develop the expression for the optimal capacity expansion, we begin by substituting the above expression for the optimal lead time, $p = \frac{1}{b}\left[a - cl + ds + e(1-s)\pi - \lambda\right]$ and $e^{-(\mu-\lambda)l^*} = 1/x$ into the objective function (9) and obtain the following objective function.

$$\prod(\lambda, z) = \left[a - \frac{c\log x}{(\mu(1+z)-\lambda)} + \frac{d\pi}{x} - \lambda - bm\right]\frac{\lambda}{b} - \frac{1}{x}\frac{\pi\lambda}{\mu(1+z)-\lambda} - k\mu z \tag{20}$$

We first differentiate $\prod(\lambda, z)$ with respect to z and obtain the first order condition:

$$\frac{\partial\prod(\lambda,z)}{\partial z} = \frac{G\lambda\mu}{b(\mu(1+z)-\lambda)^2} - k\mu = 0 \Leftrightarrow z = \frac{1}{\mu}\left(\lambda + \sqrt{\frac{G\lambda}{bk}}\right) - 1 \tag{21}$$

The above condition gives the optimal capacity expansion. To show that the optimal arrival rate, we substitute the above expression for z into $\prod(\lambda, z)$ to obtain:

$$\prod(\lambda) = \left[a - \frac{c\log x}{\sqrt{G\lambda/bk}} + \frac{d\pi}{x} - \lambda - bm\right]\frac{\lambda}{b} - \frac{1}{x}\frac{\pi\lambda}{\sqrt{G\lambda/bk}} - k\left(\lambda + \sqrt{\frac{G\lambda}{bk}}\right) + k\mu \tag{22}$$

Differencing $\prod(\lambda)$ we obtain

$$\prod{}'(\lambda) = \frac{a}{b} - m - k - \frac{2\lambda}{b} - \sqrt{\frac{G\lambda}{bk}} = 0 \Leftrightarrow (a - b(m+c) - 2\lambda)\sqrt{\lambda} - \sqrt{Gbk} = 0 \tag{23}$$

$$\prod{}''(\lambda) = -\frac{2}{b} - \frac{1}{2}\sqrt{\frac{Gk}{b}}\lambda^{-3/2} < 0 \tag{24}$$

Therefore $\prod(\lambda)$ is concave with respect to λ. Thus, we calculate a optimal solution, λ^* to maximize the $\prod(\lambda)$, and the optimal capacity expansion, z^* is given by $z^* = \frac{1}{\mu}\left(\lambda^* + \sqrt{\frac{G\lambda^*}{bk}}\right) - 1$.

To find the optimal price, we note that $\lambda^* = a - bp^* - cl^* + ds - b\pi/x$ and solve for p^*. □

Using the above optimality conditions, we can develop an iterative algorithm to obtain the optimal solution. The algorithm for the capacity expansion case, denoted as SRA-RAND_CE, proceeds as follows.

Modified RAND algorithm for capacity expansion case (SRA-RAND_CE)

Step 1. For given *n*, divide [0, $\frac{c}{b(1-s)}$] into *n* equally spaced values of π_j : (π_1, ⋯, π_j, ⋯, π_n). Set *j* = 0.

Step 2. Set *j* = *j*+1 and r = 0. Let $\pi_j(r) = \pi_j$.

Step 3. Set *r* = *r*+1.

Step 4. For a given value of $\pi_j(r-1)$, compute p_j, l_j and z_j.

Set $p_j(r) = p_j$, $l_j(r) = l_j$ and $z_j(r) = z_j$.

Step 5. For given values of $p_j(r)$, $l_j(r)$ and $z_j(r)$, compute π_j.

Set $\pi_j(r) = \pi_j$.

Step 6. If $\pi_j(r) \neq \pi_j(r-1)$, go to Step 3.

Otherwise, set $p_j^*(r) = p_j(r)$, $l_j^*(r) = l_j(r)$, $z_j^*(r) = z_j(r)$ and $\pi_j^*(r) = \pi_j(r)$.

Compute $\prod_j$ for this $[p_j^*(r), l_j^*(r), z_j^*(r), \pi_j^*(r)]$.

Step 8. If $j \neq m$, go to Step 2.

Otherwise, stop and select $[p_j^*(r), l_j^*(r), z_j^*(r), \pi_j^*(r)]$ with the maximum $\prod_j$.

4. COMPUTATIONAL EXPERIMENTS

4.1 Effect of service level

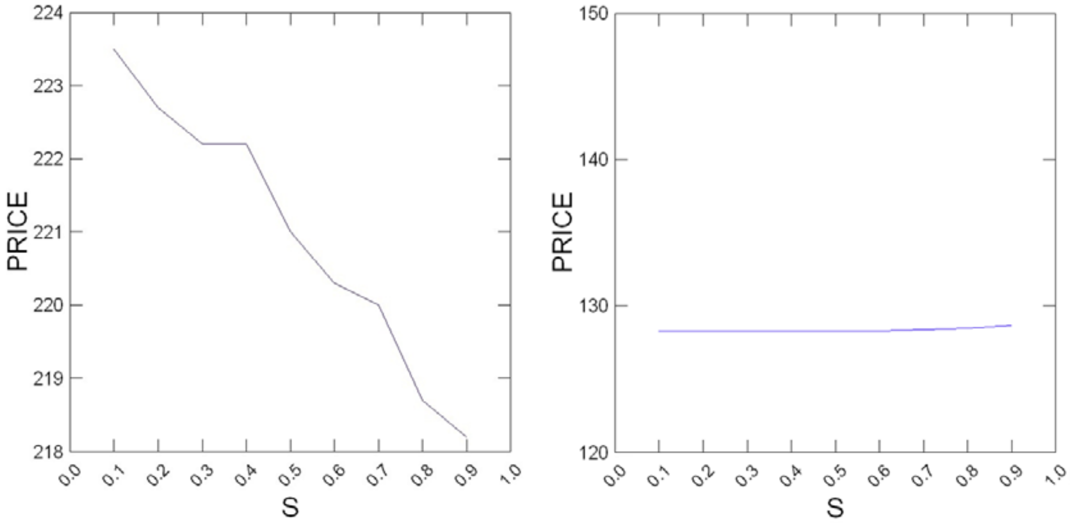
In this section, we examine the effects of service level on the optimal decision. Since the analysis is not tractable, we perform computational experiments and identify the impact through graphical means. In this study, we assume the two cases: (1) small capacity case ($\mu = a/10$) and (2) large capacity case ($\mu = a$). Since demand for the product is larger than supply capacity in small capacity case, the small capacity case represents the sellers' market while the large capacity case represents the buyers' market because a supplier has over-capacity compared with demand. The values of the parameters used in this analysis are as follows:

Table 3. The parameter values for computational experiments

parameter	value	parameter	value
a	100	m	5
b	4	μ	10, 100
c	6	s	(0, 1)
d	2		

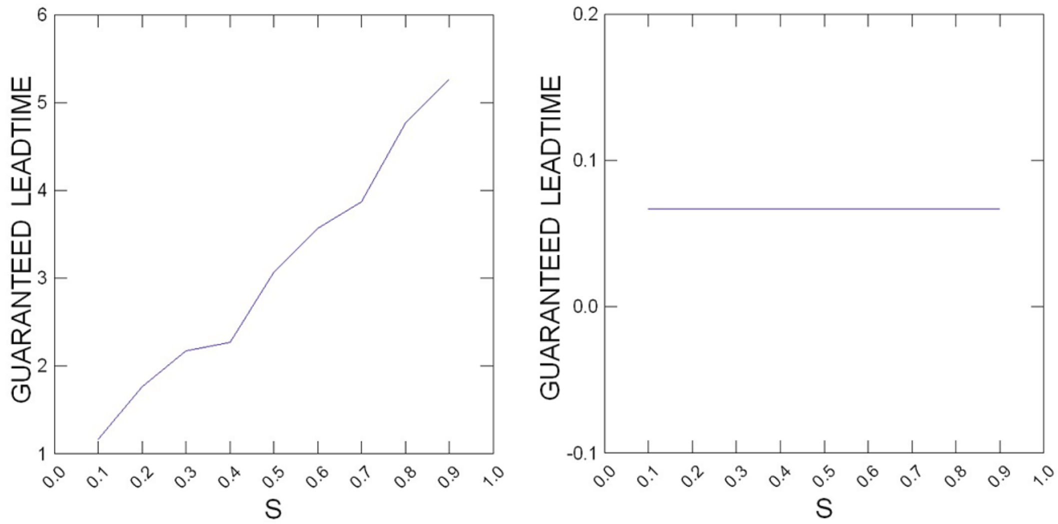
In sellers' market, the supplier determines the price while the customer determines the price in buys' market. For this reason, , the price for the product in sellers' market is larger than that in buyers' market, as shown in Figure 1. As shown in Figure 2 (a), the optimal guaranteed lead time increases as service level increases in sellers' market, so that the price for the product decreases as service level increases in sellers' market.

As shown in Figure 2, the value of optimal guaranteed lead time in sellers' market is larger than that in buyers' market, and the optimal guaranteed lead time in buyers' market close to zero. This is because supplier has sufficient capacity to meet the demand in buyers' market.



(a) $\mu = 10$ (b) $\mu = 100$

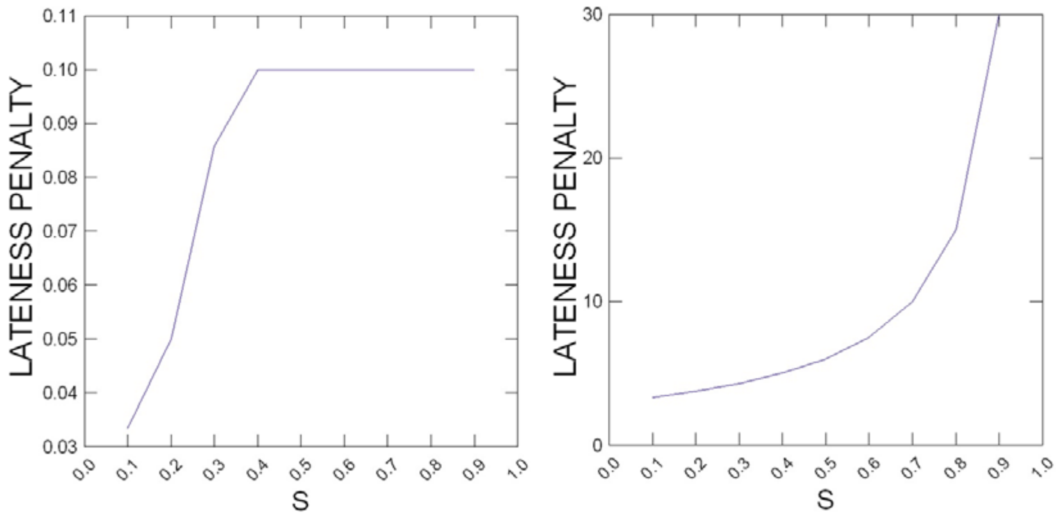
Figure 1. Optimal price, p , versus service level, s



(a) $\mu = 10$ (b) $\mu = 100$

Figure 2. Optimal guaranteed lead time, l , versus service level, s

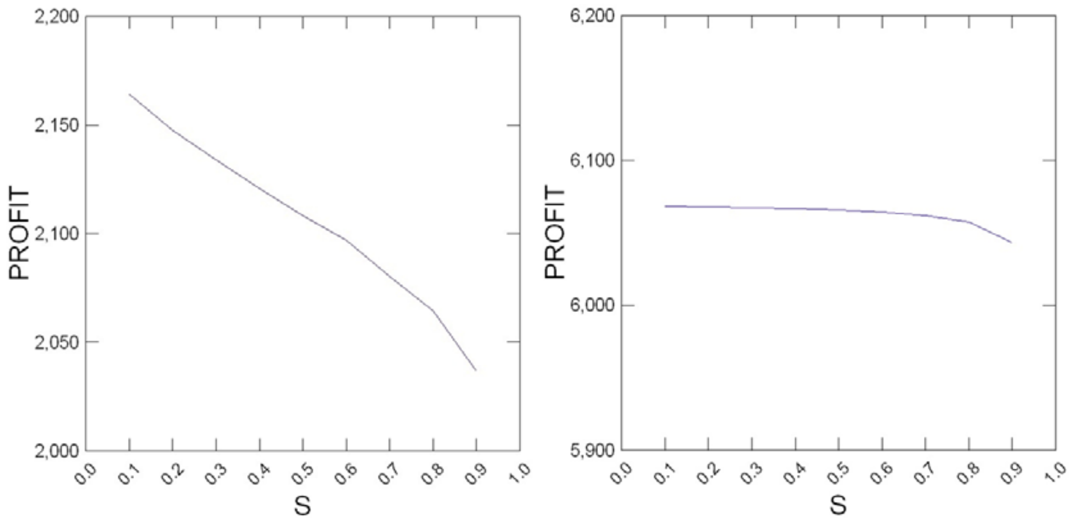
From Fig. 3 , we note that the lateness penalty close to zero in sellers' market while the lateness penalty is large in buyers' market. This is because the supplier does not need to offer lateness penalty to customer since the demand exceed the supply capacity in sellers' market. However, the supplier offers the lateness penalty to attract the customer in sellers' market.



(a) μ = 10 (b) μ = 100

Figure 3. Optimal lateness penalty, π , versus service level, s

From Fig. 4 (a), we note that the profit per unit in sellers' market is larger than that in buyers' market. Explained ahead, the supplier determines the price in sellers' market while the customer determines the price in buys' market. For this reason, the profit per unit in sellers' market is larger than that in buyers' market.



(a) μ = 10 (b) μ = 100

Figure 4. Profit versus service level, s

4. 2 Sensitivity analysis of parameters

We now examine the managerial implications of our results based on sensitivity analysis with respect to the model parameters. Table 4 summarized the parameters and their respective values employed in this study, and Table 5 presents how changes in various model parameters affect decision variables.

Table 4. The parameter values for computational experiments

parameter	value	parameter	value
a	(1,100)	d	(1,10)
b	(1,10)	m	(1,10)
c	(1,10)		

Table 5. The sensitivity of the results with respect to the model parameters

parameters	***Optimal price***		***Optimal lead time***		***Optimal lateness penalty***	
	Non-binding	***binding***	**Non-binding**	***binding***	**Non-binding**	***binding***
Maximum arrival rate, *a*	decreases	decreases	increases	increases	no impact	decreases
Price sensitivity, *b*	increases	increases	decreases	decreases	decreases	increases
Lead time sensitivity, *c*	decreases	decreases	decreases	decreases	increases	increases
Penalty cost sensitivity, *d*	decreases	decreases	increases	increases	no impact	decreases
Unit variable cost, *m*	increases	increases	decreases	decreases	no impact	increases

5. CONCLUSION

This paper studies a price, guaranteed lead time and lateness penalty decision for a supplier that offers a fixed guaranteed lead time and pays the lateness penalty to customer and presents a framework for determining the optimal price, guaranteed lead time and lateness penalty. We first develop a mathematical model for a given capacity to determine the optimal price, guaranteed lead time and lateness penalty to maximize the total profit. We then consider the case where it is also possible for the supplier to increase the capacity marginally. We also have demonstrated the impact of parameter values on the optimal decision.

An interesting extension of this research seems possible. In this paper, we assumed that lateness penalty is constant regardless of the length of lateness. Thus, we can extend our results to the case where the lateness penalty is a function of the length of lateness.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2010-0015413).

6. REFERENCES

Allon G. and Federgruen A. (2008) Competition in service industries with segmented markets. *Working paper*, Graduate School of Business, Columbia University, NY, USA.

Blackburn J. D. (1991) *Time-Based Competition: The Next Battleground in American Manufacturing*. Business One Irwin, Homewood, IL, USA.

Boyaci T. and Ray S. (2003) Product differentiation and capacity selection cost interaction in time and price sensitive markets. *Manufacturing and Service Operations Management 5(1):* 18–36.

Boyaci T. and Ray S. (2006) The impact of capacity costs on product differentiation in delivery time, delivery reliability and price. *Production and Operations Management 15(2):* 179–197.

Hill A. V. and Khosla I. S. (1992) Models for optimal lead time reduction. *Production and Operations Management 1(2):* 185–197.

Hopp W. J. and Spearman M. L. (2000) *Factory Physics*. McGraw-Hill Higher Education, USA, 2nd edition.

Hum S. H. and Sim H. H. (1996) Time-based competition: Literature review and implications for modeling. International *Journal of Operations & Production Management 16(1):* 75–90.

Kaspi M. and Rosenblatt M. (1991) On the economic ordering quantity for jointly replenished items. *International Journal of Production Research 29:* 107-114.

Palaka K. S., Erlebacher S. and Kropp D. H. (1998) Lead time setting, capacity utilization, and pricing decisions under lead time demand*. IIE Transactions 30(2):* 51–163.

Pekgun P., Griffin P. M. and Keskinocak P. (2006) Centralized vs. decentralized competition for price and lead-time sensitive demand. *Working paper*, School of Systems and Industrial Engineering, Georgia Institute of Technology, Atlanta, GA, USA.

Ray S. and Jewkes E. M. (2004) Customer lead time management when both demand and price are lead time sensitive. *European Journal of Operational Research 153(3):* 769–781.

Reddy S. K., Holak S. L. and Bhat S. (1994) To extend or not to extend: success determinants of line extensions. *Journal of Marketing Research 31(2):* 243-262.

So K. C. (2000) Price and time competition for service delivery. *Manufacturing and Service Operations Management 2(4):* 392–409.

So K. C. and Song J. S. (1998) Price, delivery time guarantees and capacity selection. *European Journal of Operational Research 111(1):* 28–49.

Stalk G. and Hout T. M. (1990) *Competing Against Time*. The Free Press, NY, USA.

Stevenson W. J. (1999) *Production and Operations Management*. Irwin McGraw-Hill, USA.

Suri R. (1998) *Quick Response Manufacturing: A Companywide Approach to Reducing Lead Times.* Productivity Press, USA.

Tsay A. and Agrawal A. N. (2000) Channel dynamics under price and service competition. *Manufacturing and Service Operations Management 2(4):* 372–391.

Van Beek P. and Van Putter C. (1987) OR contributions to flexibility improvement in production/inve ntory systems. *European Journal of Operational Research 31:* 52–60

ANALYTICAL AND SIMULATION APPROACHES OF SHIP TRAFFIC MODELING

Branislav Dragović[1*], Nam-Kyu Park[2], Romeo Meštrović[1]

[1*]Maritime Faculty, Maritime Transport & Traffic Division
University of Montenegro
Dobrota 36, 85330 Kotor
Montenegro
e-mails: branod@ac.me; bdragovic@t-com.me

[2]Department of Logistics Management
Tongmyong University
535 Yongdang-dong, Nam-gu, Busan
Republic of Korea
e-mail: nkpark@tu.ac.kr

Abstract: This paper gives modeling methodology of ship traffic and performance evaluation in port. The basic approach used mathematical and simulation models. Results from both models are compared with each other. These models are developed for impact analysis of the ship traffic and patterns of arrival ships at terminal performance. The results, analysis and conclusions given here also addresses issues such as the performance criteria and the models parameters to propose an operational method that reduces average cost per ship served and increases the terminal efficiency. Both models were applied to evaluate the efficiency of Pusan East Container Terminal (PECT).

1. INTRODUCTION

This paper presents a ship traffic modeling methodology based on statistical analysis of container ship traffic data. As a seaside link at a container terminal is the large and complex system, a performance model has to be developed. The basic approach used consists of two models. The first model applies the results of the queueing model to an analytically formulated average container ship cost function in port. The aim of this function is to minimize average container ship costs in port, including the allocation planning of berths/terminal and quay cranes/berth. The second is a simulation model adapted to the problem of analyzing ship movements in port. Implementation of the presented procedure leads to the creation of a simulation algorithm that captures seaside link performance well. The analytical model has fewer inputs and requires less computational times, whilst the simulation model can handle more practical situations with more manipulated variables and less constraints.

The past 50 years show rising interest in the research of ship traffic modeling in port. Earlier research related to a general cargo port, particularly to the seaside link modeling, using queueing theory, is summarized in Noritake, 1985. Simulation models have been used extensively for the ship traffic modeling and analysis in Tiano, *et al*., 1995, Ng and Wong, 2006, Pachakis and Kiremidjian, 2003, Yamada, *et al*., 2003, Dragović *et al*., 2005 and Yeo *et al*., 2007. The integration of analytical models (optimization) and simulation has been studied in Tiano, *et al*., 1995, Yamada, *et al*., 2003, Dragović *et al*., 2006, 2007 and 2008 for the operational planning of sea-side operations at container terminals.

This paper is organized as follows. Section 2 deals with the models for the analysis of ship movement in a port by using analytical approaches. Also, this section is concerned with the evaluation of functional estimation models in container port. After a preliminary introduction dedicated to the essential properties and main features of the traffic model, a discussion of the possible use of such a model for simulation purposes is developed in Section 3. In Section 4, numerical results and computational experiments are reported to evaluate the efficiency of the models for PECT. Section 5 gives concluding remarks.

2. MODELING

The modeling and analyzing of ship traffic in port consists of setting up mathematical models and equations for describe certain stages in the functioning of the system, which indicates that each symbol has the following meaning: λ – average ship arrival rate in ships/hour; μ – average ship service rate in ships/hour; n_b – number of berths per terminal; n_c – number of quay cranes (QCs) per berths; n_s – number of ships present in port; t_w – average waiting time in hours/ship; t_s – average service time in hours/ship; t_{ws} – average time that ships spend in port in hours/ship; t_{du} – berthing/unberthing time in hours; n_{con} – number of containers loading/unloading per ship; r_{con} – QC move time in hours/container; t_c – ships' loading/unloading time in hours/ship; k_c – QC interference exponent and θ – ship traffic intensity.

The average service time, $t_s = 1/\mu$, where $\mu = (t_c + t_{du})^{-1}$, includes ships loading/unloading time t_c, in hours per container ship, expressed as $t_c = (n_{con} \cdot r_{con})/(n_c)^{k_c}$ where $k_c = (\ln(n_{con} r_{con}) - \ln(t_c))/\ln(n_c)$. It follows that

$$n_c = \left(\frac{\lambda n_{con} r_{con}}{\theta - \lambda t_{du}} \right)^{1/k_c} \tag{1}$$

Further, it can be shown that

$$t_{ws} = t_w + t_s, \text{ where } t_w(\theta) = \frac{\theta^{n_b}}{(n_b - 1)!\,\mu(n_b - \theta)^2 ((\sum_{n_s=0}^{n_b-1} (\theta^{n_s} / n_s!)) + \theta^{n_b}(n_b - \theta)\mu} \tag{2}$$

for the ($M/M/n_b$) model. Accordingly, this parameter with the notation θ, is equal $\theta = \lambda \cdot t_s = \lambda/\mu$.

In this study, above mentioned formulae has been adapted concerning the t_w (Noritake, 1985). Accordingly with it, when the ships service time has an Erlang distribution with k phases, the following equations are obtained

$$t_{ws} = t_w V_c + t_s \tag{3}$$

for the ($M/E_k/n_b$) model, where $V_c = \frac{1}{2}\left(\frac{1}{k} + 1\right)$ - the coefficient of variation of ships service time distribution and k is the number of phases of an Erlang distribution;

$$t_{ws} = t_w \cdot \left[\frac{1}{2}\left(\frac{1}{k} + 1\right) + \left(1 - \frac{1}{k}\right)\left(1 - \frac{\theta}{n_b}\right)(n_b - 1)\frac{(4 + 5n_b)^{\frac{1}{2}} - 2}{32\theta} \right] + \frac{1}{\mu} \tag{4}$$

for the ($M/E_k/n_b$) model.

2.1 Ship traffic

A queueing theory model for analyzing traffic of $N \geq 1$ ships in port is proposed at interval $(0, T)$. Suppose that interarrival times of ships are random variables determined by the distribution function $A(t)$ with $A(+0) < 1$, and the expectation value $\alpha = \int_0^\infty t dA(t) < +\infty$. The function $A(t)$ is used below to determine the probability distribution for the number $p_i(t)$, $i = 0,1,...,N$ of ships that arrive in port at an interval $(0, t)$, so that $\sum_{i=0}^{N} p_i(t) = 1$. Such introduced random flow, determined for $0 \leq t < +\propto$ is called limited recurrent flow. The probability $p_i(t)$ can be determined (Cox and Smith, 1961) as

$$p_i(t) = A_*^{(i)}(t) - A_*^{(i+1)}(t), \quad i = 0,1,...,N-1;$$
$$p_N(t) = 1 - \sum_{i=0}^{N-1} \left(A_*^{(i)}(t) - A_*^{(i+1)}(t)\right) = A_*^{(N)}(t), \tag{5}$$

where $A_*^{(i)}(t)$ is a i'th convolution of the function $A(t)$ with itself; $A_*^{(0)}(t) \equiv 1$.

Suppose that ships arrive according to a limited Poisson process, so that

$$A(t) = \begin{cases} 1 - e^{-\lambda t} & for \quad t \geq 0 \\ 0 & for \quad t < 0 \end{cases} \tag{6}$$

Then Eq. (5) (Cox and Smith, 1961) implies

$$p_i(t) = \frac{(\lambda t)^i}{i!} e^{-\lambda t}, \ \ 0 \le i < N \tag{7}$$

$$p_N(t) = 1 - \sum_{i=0}^{N-1} \frac{(\lambda t)^i}{i!} e^{-\lambda t} \tag{8}$$

It is well known that a flow composed from finite number $m > 1$ of simple mutually independent flows with rates $\lambda_1, \dots, \lambda_m$, is also a simple flow with rate $\lambda_1 + \dots + \lambda_m$. The same property is also true for the limited Poisson process described above.

Suppose that on the interval $(0,T)$ is planed arrival of $N_1 + N_2 + N_3$ ships, where the number N_i related to i'th flow, $1 \le i \le 3$, and $N_1 \le N_2 \le N_3$. Next suppose that mean interarrival time of ships in i'th flow, $1 \le i \le 3$, is λ_i^{-1}. Denote by $p_n(t)$, $0 \le n \le N_1 + N_2 + N_3$, the probability $p_n(t)$ at the interval $(0,t)$ arrive n ships from described flow. Then since the above three flows are mutually independent, we have

$$p_n(t) = \sum_{\substack{n_1+n_2+n_3=n \\ 0 \le n_i \le N_i \\ 1 \le i \le 3}} p_{n_1}^{(1)}(t) p_{n_2}^{(2)}(t) p_{n_3}^{(3)}(t), \tag{9}$$

where the $p_{n_i}^{i}(t)$, $1 \le i \le 3$, can be computed by the Eqs. (7) and (8) by changing λ with λ_i and N with N_i.

In order to determine the sum on the right hand side of the Eq. (9), we consider the following cases:

Case 1, $n_i \ne N_i$ for all $1 \le i \le 3$. Then by Eq. (7) the corresponding term in the sum given by Eq. (9) is

$$\prod_{j=1}^{3} \frac{(\lambda_j t)^{n_j}}{n_j!} e^{-\lambda_j t} = \left(\prod_{j=1}^{3} \frac{(\lambda_j)^{n_j}}{n_j!} \right) t^n e^{-\sum_{j=1}^{3} \lambda_j t} \tag{10}$$

Case 2, $n_i = N_i$ for exactly one i, $1 \le i \le 3$. Then for such an i by Eqs. (7) and (8) we have the corresponding term in the sum expressed by Eq. (9) is

$$\left(1 - \sum_{k=0}^{N_i - 1} \frac{(\lambda_i t)^k}{k!} e^{-\lambda_i t} \right) \prod_{\substack{j=1 \\ j \ne i}}^{3} \frac{(\lambda_j t)^{n_j}}{n_j!} e^{-\lambda_j t} = \left(1 - \sum_{k=0}^{N_i - 1} \frac{(\lambda_i t)^k}{k!} e^{-\lambda_i t} \right) \left(\prod_{\substack{j=1 \\ j \ne i}}^{3} \frac{\lambda_j^{\,n_j}}{n_j!} \right) t^{\sum_{j=1}^{3} n_j - n_i} e^{-\sum_{j=1}^{3} \lambda_j t - \lambda_i t} \tag{11}$$

Because of $\sum_{j=1}^{3} n_j - n_i = n - n_i = n - N_i$, the above sum is

$$\left(1 - \sum_{k=0}^{N_i - 1} \frac{(\lambda_i t)^k}{k!} e^{-\lambda_i t} \right) \left(\prod_{\substack{j=1 \\ j \ne i}}^{3} \frac{\lambda_j^{\,n_j}}{n_j!} \right) t^{n - N_i} e^{-\sum_{j=1}^{3} \lambda_j t - \lambda_i t}. \tag{12}$$

Case 3, $n_i = N_i$ and $n_l = N_l$ for exactly two values *i* and *l* with $1 \le i < l \le 3$. Then for such i and *l* by Eqs. (7) and (8) we obtain that the corresponding term in the sum given by Eq. (9) is

$$\left(1 - \sum_{k=0}^{N_i - 1} \frac{(\lambda_i t)^k}{k!} e^{-\lambda_i t} \right) \left(1 - \sum_{k=0}^{N_l - 1} \frac{(\lambda_l t)^k}{k!} e^{-\lambda_l t} \right) \frac{(\lambda_r t)^{n_r}}{n_r!} e^{-\lambda_r t} \tag{13}$$

where $1 \le i < l \le 3$ and *r* is the element from $\{1, 2, 3\}$ distinct of i and l.

Case 4, $n_i = N_i$ for all $1 \le i \le 3$. Then $n = N_1 + N_2 + N_3$ and by Eq. (8) we obtain that the corresponding term in the sum expressed by Eq. (9) is

$$\prod_{i=1}^{3} \left(1 - \sum_{k=0}^{N_i - 1} \frac{(\lambda_i t)^k}{k!} e^{-\lambda_i t} \right) \tag{14}$$

Finally, it is obvious that $p_n(t)$ determined by Eq. (9), can be expressed as a sum of three subsums whose general terms are given by Eqs. (10), (12), (13) and (14), respectively. Recall that in the whole sum given by Eq. (9) the second mentioned subsum must be written for any i with $1 \le i \le 3$, and the third mentioned subsum must be written for any pair i, l with $1 \le i < l \le 3$.

2.2 Ship operations

Model elements of the container terminal can be separated into following group: berth cost in \$ per hour, $c_1 = n_b c_{n_b}$; QC cost in \$ per hour, $c_2 = n_b n_c c_{n_c}$; storage yards cost in \$ per hour, $c_3 = \theta\mu n_{con} t_{t_{con}} a_{con_{cy}} c_{cy}$; transportation cost by yard transport equipment between quayside and storage yard in \$ per hour, $c_4 = \theta\mu n_c t_c n_{cyc} c_t$; labor cost for QC gangs in \$ per hour, $c_5 = \theta\mu n_c t_l c_l$; ships cost in port in \$ per hour, $c_6 = \theta\mu t_{ws} c_s$; and containers cost and its contents in \$ per hour, $c_7 = \theta\mu t_{ws} n_{r_{con}} c_w$. The total cost function, would be concerned as $TC = \sum_{i=1}^{7} c_i$.

It is necessary to know that only the *TC* computes the number of berths/terminal and QCs/berth that would satisfy the basic premise that the service port cost plus the cost of ships in port should be at a minimum. This function was introduced by (Sch. and Shar., 1985). We point out that their solutions may not be as good as ours because we have simulation approach to determine key parameters t_w, t_s, λ, μ, θ and especially k_c. Therefore, to find the optimal solution, their function can be obtained as

$$TC = f(\theta) = n_b\left(c_{n_b} + n_c c_{n_c}\right) + \theta\mu\left(n_{con} t_{t_{con}} a_{con_{cy}} c_{cy} + n_c t_l\left(c_l + n_{cyc} c_t\right) + t_{ws}(\theta)\left(c_s + n_{r_{con}} c_w\right)\right) \tag{15}$$

where *TC* - total port system costs in \$/hour; c_{n_b} - hourly berth cost in \$, ($c_{n_{b1}}$ - the initial berth cost, i - interest rate, n_y - economic lifetime in years, $c_{n_{bm}}$ - annual maintenance cost per berth), $c_{n_b} = \left(c_{n_{b1}}\left(i(1+i)^{n_y} / \left((1+i)^{n_y} - 1\right)\right) + c_{n_{bm}}\right)/(365 \cdot 24)$; c_{n_c} - QC cost in \$/QC hour; $t_{t_{con}}$ - average yard container dwell time, in hours; $a_{con_{cy}}$ - number of m^2 of storage yard per container; c_{cy} - container yard (CY) cost in \$/m^2 hour; n_{cyc} - hourly average number of cycle by yard transport equipment between quay side and CY; c_t - transportation cost between quay side and CY per cycle in \$; t_l - paid labor time in hour per gang per ship, $t_l = \max\{t_c\}$; c_l - labor cost in \$/gang hour; c_s - ship cost in port in \$/ship hour; $n_{r_{con}}$ - average payload in containers/ship; c_w - average waiting cost of a container and its contents in \$/container hour.

By substituting Eq. (1) into Eq. (15) we obtain (Dragović *et al.*, 2007 and 2008)

$$TC = f(\theta) = n_b c_{n_b} + \lambda n_{con} t_{t_{con}} a_{con_{cy}} c_{cy} + \left(\frac{\lambda n_{con} r_{con}}{\theta - \lambda t_{du}}\right)^{1/k_c}\left(n_b c_{n_c} + \lambda t_l\left(c_l + n_{cyc} c_t\right)\right) + \lambda t_{ws}(\theta)\left(c_s + n_{r_{con}} c_w\right) \tag{16}$$

where $t_{ws}(\theta)$ is defined by the Eq. (3) or Eq. (4) or by a result of simulation modeling.

From the total port cost function per average arrival rate, we can obtain the average container ship cost in \$/ship, *AC*

$$AC = \frac{f(\theta)}{\lambda} = \frac{f(\theta)}{\theta\mu} \tag{17}$$

3. SIMULATION AS MORE GENERAL MODEL

Most container terminal systems are sufficiently complex to warrant simulation analysis to determine systems performance. The GPSS/H simulation language, specifically designed for the simulation of manufacturing and queueing systems, has been used in this paper. In order to present the seaside link processes as accurate as possible the following phases need to be included into simulation model (Dragović *et al.*, 2006 and 2007): **Model structure –** seaside link is complex due to different interarrival times of ships, different dimensions of ships, multiple quays and berths, different capabilities of QCs and so on. The modeling of these systems must be divided into several segments, each of which has its own specific input parameters; **Data collection -** All input values of parameters within each segment are based on data collected in the context of this research. The main input data consists of ship interarrival times, lifts per ship, number of allocated QCs per ship call, and QC productivity. Existing input data are subsequently aggregated and analyzed so that an accurate simulation algorithm is created in order to evaluate seaside link; **Inter-arrival times of ships -** The inter-arrival time distribution is a basic input parameter that has to be assumed or inferred from observed data. The most commonly assumed distributions in literature are the exponential distribution; the negative exponential distribution or the Weibull distribution (Dragović *et al*., 2007); **Loading and unloading stage -** Accurate representation of number of lifts per ship call is one of the basic tasks of seaside link modeling procedure. It means that, in accordance with the division of ships in different classes, the distribution corresponding to those classes has to be determined; **Number of QCs per ship -** The data available on the use of QCs in seaside operations have to be considered too, as this is another significant issue in the service of ships. This is especially important as total t_{ws} depends not only on the number of lifts but also on the number of QCs allocated per ship. Different rules and relationships can be used in order to determinate adequate number of QCs per ship; **Flowchart -** After the input parameter is read, simulation starts by generating ship arrivals according to the stipulated distribution.

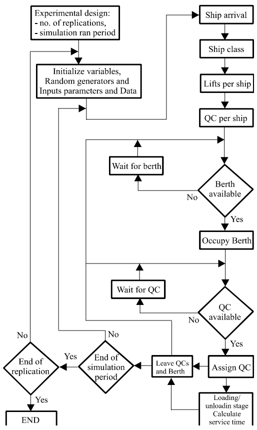


Figure 1. Flowchart for a ship arrival/departure

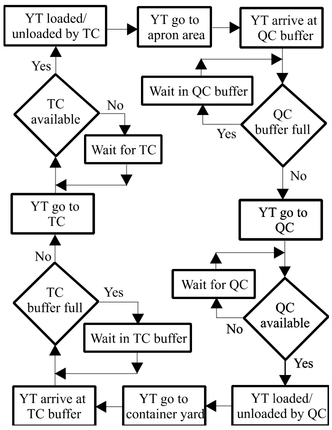


Figure 2. Flowchart of the terminal transport operations

Then, the priority of the ship is assigned depending on its size. The ship size is important for making the ship service priority strategies. For the assumed number of lifts per ship to be processed, the number of QCs to be requested is chosen from empirical distribution. This procedure is presented in the algorithms shown in Fig. 1. In order to calculate the seaside link performance, it is essential to have a through understanding of the most important elements in a port system including ship berthing/unberthing, QCs/ship, yard tractor allocation to a container and QC allocation in stacking area. As described in Fig. 2 - process flow diagram of the terminal transport operations, the scope of simulation, strategy and initial value and performance measure will have to be defined. To move containers from apron to stacking area, four tractors are provided for each QC.

4. RESULTS

To demonstrate the application of the models of the previous sections to the container terminal system, some computational results are presented for both of them. Here, we give a seaside link modeling methodology based on statistical analysis of container ship traffic data obtained from the PECT. PECT is big container terminals with a capacity of 2,075,895 twenty foot equivalent units (TEU) in 2006. There are five berths with total quay length of 1,500 m and draft around 14-15 m, Fig. 3 (Dragović *et al*., 2007 and 2008). Ships of each class can be serviced at each berth. An important part of the model implementation is the correct choice of the values of the simulation parameters.

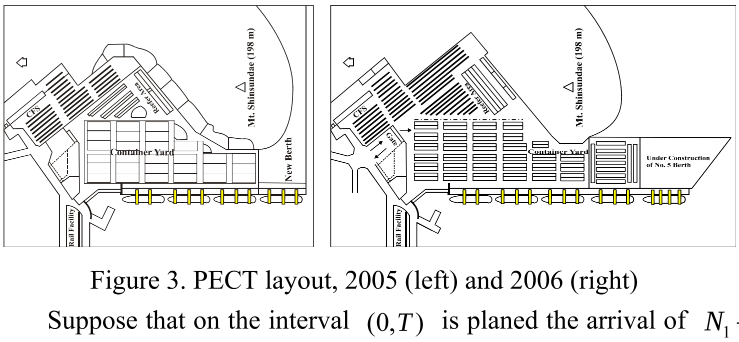


Figure 3. PECT layout, 2005 (left) and 2006 (right)

The input data for the both models are based on the actual ship arrivals at the PECT from January 1, 2005 to October 31, 2005 (Fig. 3, left) and January 1, 2006 to October 31, 2006 (Fig. 3, right), respectively (Dragović *et al*., 2007 and 2008). The ship arrival rate was 0.168 in 2005 and 0.176 ships/hour in 2006. Total throughput during the considering period was 1,704,173 TEU in 2005 and 1,703,662 TEU in 2006. The ships were categorized into the three classes according to the number of lifts, Table 1.

Suppose that on the interval $(0,T)$ is planed the arrival of $N_1+N_2+N_3$ ships, where the number N_i related to i'th flow (class), $1 \le i \le 3$, and $N_1 \le N_2 \le N_3$. Next suppose that mean inter-arrival time of ships in i'th flow, $1 \le i \le 3$, is λ_i^{-1}. Since numbers of ships in each of three ship classes given in the Table 1, we can assume that $\lambda_1 : \lambda_2 : \lambda_3 = N_1 : N_2 : N_3$. In particular, for $T = 1$ hour we obtain $\lambda_1 = 0.0399$, $\lambda_2 = 0.0685$ and $\lambda_3 = 0.0594$ for 2005. Similarly, by using the same notation for 2006 from Table 1 we have $\lambda_1 = 0.0526$, $\lambda_2 = 0.0663$ and $\lambda_3 = 0.0570$. By using these value from Eqs. (9) and (10), we can compute the probabilities $p_n(t)$ for small values of n. The analogous results may be obtained for arbitrary times less than 10 months.

Table 1. Actual ship arrivals at the PECT in 2005 and 2006

Class of ships	Number of ships		Average lifts per ship		Total lifts	
	2005	2006	2005	2006	2005	2006
< 500 lifts	292	384	313	305	91,394	117,405
501 – 1000 lifts	500	485	782	741	364,100	359,401
> 1000 lifts	433	416	1,444	1,413	625,371	587,983
All classes	1,225	1,285	882	829	1,080,865	1,064,789

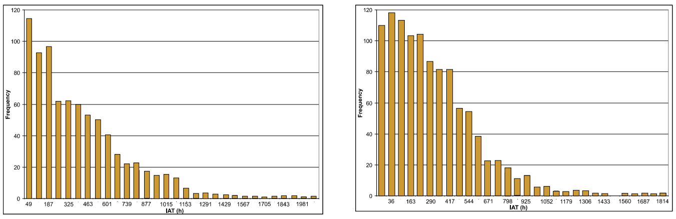


Figure 4. IATD of ships at PECT in 2005 (left) and in 2006 (right)

The interarrival time distribution (IATD) is plotted in the Fig. 4. Interestingly, even though ship arrivals of the ships are scheduled and not random, the distribution of interarrival times fitted very well the exponential distribution. Service times were calculated by using the Erlang distribution with different phases. To obtain accurate data, we have first fitted the empirical distribution of service times of ships to the appropriate theoretical distribution.

Service time distributions are given in 2005: Service distribution (SD) of first class of ships, the 4-phase Erlang distribution, (E_4); SD of second class of ships, (E_4); SD of third class of ships, (E_5) and SD of all classes of ships, (E_3). It is observed that for 2006 service time of the first ship class follows the 5-phase Erlang distribution, while the 6-phase Erlang distribution fits very well the service time of the second ship class, than 2-phase Erlang distribution fits very well the service time of the third ship class and all classes of ships follows 4-phase Erlang distribution. Goodness-of-fit was evaluated, for all tested data, by both chi-square and Kolmogorov-Smirnov tests at a 5 % significance level.

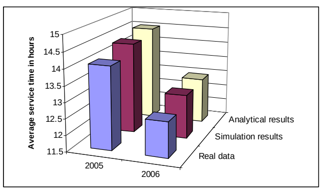
We have carried out extensive numerical work for high/low values of the PECT model characteristics. Our numerical experiments are based on different parameters of various PECT characteristics presented in Table 2.

Table 2. Input data – Terminal characteristics (Dragović *et al.*, 2007 and 2008)

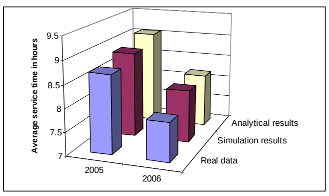
Class of ships	Input data											
	n_{con} (no. of con.)		r_{con} (hrs per con.)		t_l (hrs/gang/ship)		c_s ($/ship hrs)		n_c^*		k_c	
	2005	2006	2005	2006	2005	2006	2005	2006	2005	2006	2005	2006
First	313	305	0.05	0.05	8.80	7.1	745	739	1.7	1.9	0.80	0.90
Second	782	741	0.05	0.05	13.9	12.7	1098	1081	2.5	2.6	0.93	0.925
Third	1444	1413	0.03	0.03	20.2	17.8	1354	1365	3.1	3.3	0.97	0.965
All classes	862	829	0.04	0.04	14.3	12.9	1164	1155	2.5	2.6	0.91	0.911

n_c^* - average number of QCs assigned per ship (Real data and Simulation results); $c_{n_{b1}}$ = 62 million \$; i = .0663; n_y - 40, $c_{n_{bm}}$ = 6.2 million \$; c_{n_b} = 1215 \$; c_{n_c} = 38.8 \$/QC hour; $t_{t_{con}}$ = 188 hours; $a_{con_{cy}}$ = 63.9 m^2/container; C_{cy} = 0.000292 \$/m^2 hour; n_{cyc} = 9; c_t = 5 \$/cycle; c_l = 357 \$/gang hour; $n_{r_{con}}$ (601 for I class, 1085 for II class, 1312 for III class, 999 for all classes in 2005; and 642 for I class, 1114 for II class, 1371 for III class, 1042 for all classes in 2006); c_w =1.4 \$/container hour. To move containers from apron to CY, four tractors are provided for each QC. It takes average 10 minutes from apron to CY including handling time by transfer crane. The average distance between apron and CY is assumed to be 850 meters.

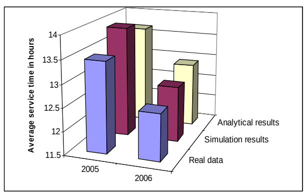
The impact of the different models is determined by comparing the key performance measures of analytical and simulation approaches to those of the real data of PECT. Fig. 5 presents t_s, while Fig. 6 shows t_w. In addition, Fig. 7 gives t_{ws}. According to this, judging from the computational results for some numerical examples of the $(M/E_k/n_b)$ – using t_w from Eq. (3) (for brevity analytical model I (AM I)) and $(M/E_k/n_b)$ – using t_w from Eq. (4) (for brevity analytical model II (AM II)) models, it can be confirmed that Eq. (3) is inclined to estimate the values of t_{ws}, i.e. t_w.



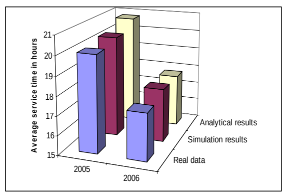
a) All classes



b) First class

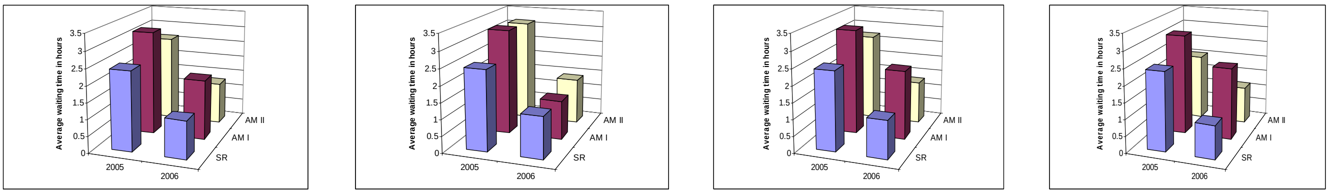


c) Second class



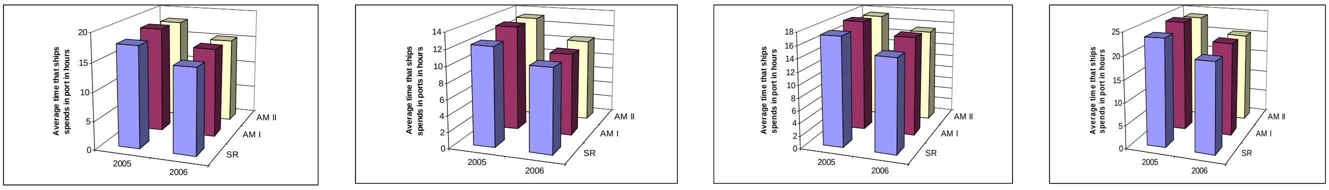
d) Third class

Figure 5. Average service time of ships in hours



a) All classes b) First class c) Second class d) Third class

Figure 6. Average waiting time of ships in hours



a) All classes b) First class c) Second class d) Third class

Figure 7. Average time that ships spend in port

The t_{ws} for simulation model (SM) is 15.036 hrs for all classes of ships in 2006. This is about 15% shorter than that of SM, 17.799 h in 2005 and about 1.5% shorter than that of AM II, 15.245 hrs in 2006. For first class of ships, the average time that ships spend in port is 10.380 hrs for AM II in 2006, about 0.6% shorter than SM, 10.441. This time is 15.232 hrs for second class of ships (AM II) in 2006, about 12% shorter than AM II in 2005. Finally, the t_{ws} for third class of ships is 19.818 hrs (SM) in 2006, about 16% shorter than SM in 2005 or 2% shorter than AM II, 20.270 h in 2006.

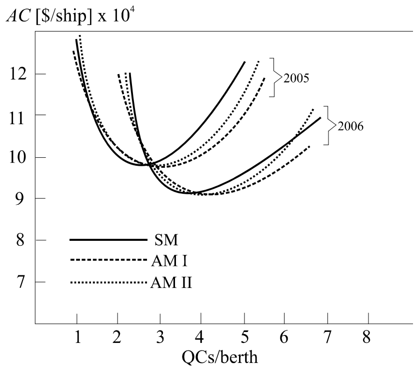


Figure 8. *AC* for various QCs/berth

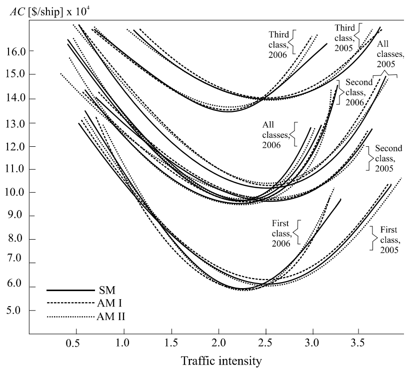


Figure 9. *AC* in function of θ

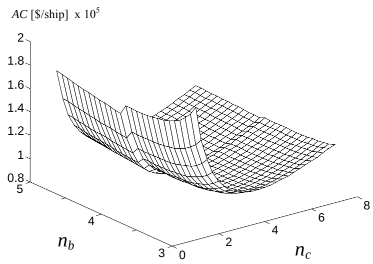


Figure 10. *AC* for various n_b and n_c

Fig. 8 presents how additional QCs reduce the *AC* for each model in relation to all classes of ships. In curve SM, the minimum cost per ship served decreases by about 7.5% in 2006 with respect to 2005. This decrease is about 7.5% in 2006 with respect to 2005 for curve AM I from Fig. 8. Finally, in curve AM II from Fig. 8, the minimum cost per ship served decreases by about 8% in relation to 2005. Figure 9 compares the *AC* of different ship classes taken by SM, AM I and AM II models at a PECT in 2005 and 2006. They graphically show the sensitivity of the *AC* to the various values of θ. In curve SM for all classes of ships in 2006, the minimum cost per ship served decreases by about 3.3% in 2006 with respect to 2005. However, the *AC* per first class of ships served decrease in 2006 by about 7% than the minimum cost in 2005, see curve AM II. This decrease for second class of ships is about 2% in 2006 with respect to the minimum cost in 2005 for curve AM II. Finally, in curve SM for third class of ships, the minimum cost per ship served decreases by about 2.6% ($138,019) than the minimum cost in 2005 ($141,697). Fig. 10 shows the optimization function *AC* of two variables n_b and n_c for constant value of θ. These results correspond to those from Fig. 9. Still, even in Fig. 10, the study offers similar results, i.e. the minimum average cost per ship served are $96,721 in relation to $96,721 from Fig. 9 – curve AM II for second class of ships. All results presented here are obtained by using the input data from Table 2. Simulation testing was than carried out by using the GPSS/H. The solution procedure for AM was programmed using the MATLAB.

5. CONCLUSIONS

Our results show that ships arrivals over time are needed as input data for the optimisation of the problem. In addition to the arrival date and ships time in port, it also generates the number of lifts per ship. On the basis on a QC productivity, this number of lifts per ship can easily be converted into the average service time of ships needed at the berth.The results presented here support the argument that the *AC* could be decreased by increasing number of QCs/ship and their productivity. The attained agreement of the results obtained by using SM with corresponding values of real parameters has also been used for validation and verification of applied AM. The correspondence between simulation and analytical results completely shows the validity to the applied AM.

These results will emphasize the effects of terminal and traffic intensity, average time that ships spend in port, numbers of QCs/berth, QC productivity and numbers of berths/terminal. These five parameters are keys to the analysis of the whole container port efficiency and achievement of economies of scale. However, major improvements in port productivity, quality of service and costs reduction can be achieved by joint optimizing these variables.

ACKNOWLEDGEMENTS

A part of this work is a contribution to the Ministry of Science and Education of Montenegro funded project No. 05-1/3-3271.

6. REFERENCES

Cox, D.R. and Smith, W. L. (1961). *Queues*, London: Methuen and Co.

Dragović B., Park N. K., Radmilović Z. and Maraš V. (2005). Simulation modeling of ship-berth link with priority service, *Maritime Economics & Lo*gistics, 7(4), 316-335.

Dragović B., Park N. K., Radmilović Z. (2006). Ship-berth link performance evaluation – simulation and analytical approaches –, *Maritime Policy & Management*, 33(3), 281-299.

Dragović B., Park N. K., Radmilović Z. (2007). Anchorage-ship-berth-yard link modeling in container port, *Proceedings of the* 11th WCTR 2007, Berkeley, CA, USA, 1-21.

Dragović B., Park N. K., Ryoo D-K., Meštrović R. (2008). Ship traffic in container port: Modeling methodology and performance evaluation, *Proceedings of First International Forum of Shipping, Port and Airport*, China, HK, 1-15.

Noritake, M. (1985). Congestion cost and pricing of seaports, *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 111(2), 354-370.

Ng, W.C. and Wong C.S. (2006). Evaluating the impact of vessel-traffic interference on container terminal capacity, *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 132(2), 76-82.

Pachakis, D. and Kiremidjian, A. S. (2003). Ship traffic modeling methodology for ports, *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 129(5), 193-202.

Schonfeld P., Sharafeldien O. (1985). Optimal berth and crane combinations in containerports, *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 111(6), 1060-1072.

Tiano, A., Cuneo, M. & Fratta, C. (1995). Ship traffic modelling and simulation, Paper in Book: *Marine Technology and Transportation*, Wessex Institute of Technology, U.K.

Yamada, T., Frazila, R.B., Yoshizawa, G., Mori, K. (2003). Optimising the handling capacity in a container terminal for investigating efficient handling systems, *J. of the Eastern Asia Society for Transportation Studies*, 5(1), 597-608.

Yeo, G-T., Roe, M. and Sang-Moon Soak, S-M. (2007). Evaluation of the marine traffic congestion of North Harbor in Busan Port, *Journal of Waterway, Port, Coastal, and Ocean Engineering***,** 133(2), 87-93.

SELECTION OF BENCHMARK TARGETS USING DEA CONSIDERING DIRECTION AND UNIT SIMILARITY

Jaehun Park[1], Hyerim Bae[1*], Sungmook Lim[2] and Loo Hay Lee[3]

[1]Department of Industrial Engineering
Pusan National University
Busan, Korea
e-mail: pjh3479@pusan.ac.kr, *hrbae@pusan.ac.kr

[2]Division of Business Administration
Korea University
Jochiwon, Korea
e-mail: sungmook@korea.ac.kr

[3]Department of Industrial & System Engineering
National University of Singapore
Singapore
e-mail: iseleelh@nus.edu.sg

Abstract: Benchmarking requires an effective methodology for finding the best performer, which entails an evaluation of the relative efficiencies of competitors in terms of multiple input and output factors. To identify the best performer, Data Envelopment Analysis (DEA) has been popularly used. However, the conventional DEA has some deficiencies with respect to its use for benchmarking. First, benchmark targets derived from DEA analysis might be a hypothetical Decision Making Unit (DMU) that does not actually exist. Second, the reference set of an inefficient DMU often has multiple efficient DMUs. Third, it might be quite impossible for an inefficient DMU to achieve its target's efficiency in a single step, especially when the target is far removed from the DMU. To overcome these deficiencies of conventional DEA, we propose a new stepwise benchmarking method using DEA, which enables inefficient DMUs to select the more appropriate benchmarking DMU based on the direction and similarity.

1. INTRODUCTION

Benchmark target selection has been recognized important factors for the inefficient organizations to improve its efficiencies. And several studies on benchmarking target selection have been conducted in various fields such as public administration (Ammons, 2002), production and design literature (Grupp, 1990), and business management (Tata *et al*, 2000). In general, a benchmarking process consists of three steps. The first step is identifying a company that is acknowledged as the best performer and the second is setting benchmarking goals and the third is implementing the best practices (Donthu *et al.*, 2005). During the benchmarking procedure, an effective methodology for finding the best performer is inevitably required, which entails evaluation of the relative efficiencies of competitors according to multiple input and output factors. Therefore, identifying the best performer can be considered to be the most important activity in the benchmarking process.

To identity a best performer, Data Envelopment Analysis (DEA), a methodology for measuring the relative efficiencies among homogeneous Decision-Making Units (DMUs), has been used (Ross and Droge, 2002). DEA accomplishes this task by means of multiple inputs and outputs, yielding a reference target for an inefficient DMU along with the corresponding efficiency gap (the degree to which that DMU must be improved so as to be rendered efficient). However, the general use of DEA has certain limitations in aspects of benchmarking which are underlined below;

1) The reference target might be a hypothetical DMU that does not actually exist (it is difficult and indeed unrealistic to learn from such a DMU).
2) An efficient DMU may have multiple benchmarks that it should utilize as possible targets for improvement. This may pose a practical problem because it becomes difficult to simultaneously identify 'best practices' from a variety of sources in improving the performance of an inefficient DMU.
3) For a large sample data, it is quite impossible for an inefficient DMU to achieve its target's efficiency in a single step, especially when that DMU is far from the benchmarking target DMU on efficient frontier.

In order to overcome these problems, various benchmarking method have been undertaken. The research objective of the existing work is categorized into two area; benchmarking target selection and stepwise improvement.

Typically, the work for benchmarking target selection includes models that enhance DEA flexibility in estimating targets for relative inefficient DMUs. Thanassoulis and Dyson (1989) developed a DEA model that sets efficient benchmarking targets for inefficient DMUs. They point out that it would be desirable to take not only the nature of the controllability of their inputs and outputs but also of the priorities of improving individual inputs and outputs into account. Bogetoft and Hougaard (1999) introduced the potential improvement index. That efficiency index will guide the selection of reference plans. Gonzales and Alvarez (2001) developed a model based on the input-contraction method that computes the sum of input contractions required to reach the efficient subset of the production frontier.

The work for the stepwise improvement provides a stepwise path for improving the efficiency of each inefficient DMU. Joe (2003) proposed a stratification method by iteratively generating the efficient frontiers. Alirezaee and Afsharian (2007) proposed a layered efficiency evaluation model that provides a strategy by which an inefficient DMU can move toward a better layer. However, this model lacks information on how to choose the reference DMU on each layer. Shaneth *et al.* (2009) proposed a proximity-based target selection method to provide, using Self-Organizing Map (SOM) and Reinforcement Learning, the optimal path to the most efficient frontier DMU by means of a minimal efficiency gap. This method considers not the reference set of inefficient DMUs, but practical target DMUs based on the similarity of input patterns for the benchmarking path. Park *et al.* (2010) proposed a stratification benchmarking path method that, with the benchmarking path provided, can reduce the distance between an inefficient DMU and an efficient DMU by clustering DMUs based on input-similarity.

The existing works addressed here can be considered to more realistic and more effective than conventional DEA approaches, because these overcome the limitations of conventional DEA in aspects of benchmarking and propose the stepwise benchmarking DMU for each inefficient DMU based on the efficiency. There are many strategies and methods, both actual and potential for selection of the next benchmarking DMU. However, the existing stepwise benchmarking methods are still limited in that they consider only the efficiencies and do not consider various strategies such as direction and similarity for rendering inefficient DMUs efficient.

In this paper, we propose a new stepwise benchmarking method using DEA, which enables inefficient DMU to select the next step benchmarking DMU based on the improving direction and unit similarity to overcome the above mentioned deficiencies of conventional DEA and stepwise benchmarking methods. Improving direction is employed to select the closest efficient DMU for the next step benchmarking DMU, and unit similarity is utilized to select similar input-factor DMUs with Self-Organizing Map (SOM). Additionally, an integrated benchmarking method combining improving direction and unit similarity for the selection of the next stepwise benchmarking DMUs using DEA is discussed. As an application of the proposed method, benchmarking of East Asia container terminal has been conducted. The structure of this paper is organized as follows. Section 2 provides an overview of DEA and SOM. Section 3 discusses the proposed method, and section 4 details our empirical study. Finally Section 5 summarizes our work.

2. RESEARCH BACKGROUND

2.1 Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) is a linear-programming methodology that evaluates the relative efficiencies of DMUs using a set of inputs to produce a set of outputs (Joe, 2003). The mathematical model of DEA is represented by Equation (1) below.

$$\max \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}}$$
$$s.t. \quad \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1; \quad j = 1,\dots,n \qquad (1)$$
$$u_r, v_i \ge 0; r = 1,\dots,s; i = 1,\dots,m$$

This (Equation (1)), a CCR model, is a basic DEA model initially developed by Charnes, Cooper & Rhodes (1978). Here, u_r is the weight given to the r-th output, v_i is the weight given to the i-th input, n is the number of DMUs, s is the number of outputs, m is the number of inputs, k is the DMU being measured, y_{rj} is the amount of the r-th output produced by DMU j, and x_{ij} is the amount of the i-th input produced by DMU j. The DEA model can be divided into input-oriented and output-oriented versions, according to the rationale for conducting DEA. The input-oriented model minimizes inputs with the given outputs, whereas the output-oriented model maximizes outputs with the given inputs. The fractional model shown as (1) can be converted to a linear model. For more details on model development refer to

Charnes *et al*. (1978).

DEA can be a beneficial tool to improve performance through efficiency evaluation and benchmarking, specifically by suggesting a reference set, which is a set of corresponding efficient units that can be utilized as a benchmark for improvement. The reference set can be obtained by dual model, as shown in (2).

$$\min \theta - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right)$$

$$s.t.\ \sum_{j=1}^{n} \lambda_j x_{ij} - \theta x_{ik} + s_i^- \le 0,\ \ (i = 1,2,...,m), \tag{2}$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - y_{rk} - s_r^+ \ge 0,\ \ (r = 1,2,...,s),$$

$$\lambda_j,\ s_i^-,\ s_r^+ \ge 0, \qquad (j = 1,2,...,n)$$

In model (2), θ is the efficiency score, λ_j is the dual variable, and ε is a non-Archimedean infinitesimal. By solving model (2), we can identify a composite DMU (a linear combination of DMUs) that utilizes less input than the test DMU while maintaining at least the same output levels. The optimal values of the dual variable λ_j, are the coefficients for this linear combination of units. The set of units involved in the construction of the composite DMU can be utilized as a benchmark for improvement of the inefficient test DMU. If a DMU is given an efficiency score of '1', it is considered to be efficient; an efficiency score less than '1' indicates inefficiency.

3. THE PROPOSED METHOD

In this section, we address the drawbacks of conventional DEA and stratification method in aspects of benchmarking, and to remedy these drawbacks, we propose a new stepwise benchmarking method, which enables inefficient DMU to select the benchmarking DMUs, based on improving direction and unit similarity.

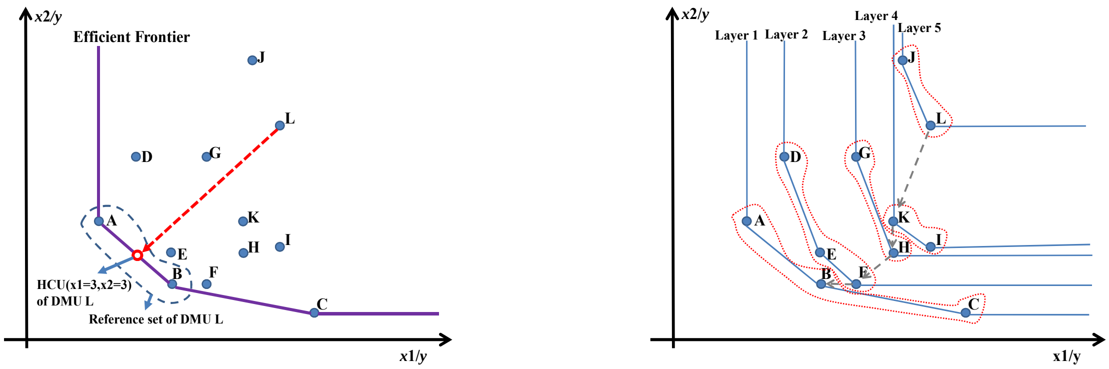
Let's consider the following supermarket example, originally introduced in Copper *et al*. (2006) and more DMUs are added for easy explanation of the procedure outlined in this paper. Table 1 consists of twelve DMUs, and each DMU consumes two inputs and yields one output. The two inputs and the one output are the number of employees (unit: 10), the floor area (unit: 1000 m^2) and the sales (unit: \$100,000), respectively.

Table 1. Supermarket Example

Store		A	B	C	D	E	F	G	H	I	J	K	L
Employee	x_1	2	4	8	3	4	5	5	6	7	6	6	7
Floor area	x_2	4	2	1	6	3	2	6	3	3	9	4	7
Sales	y	1	1	1	1	1	1	1	1	1	1	1	1

When we apply the input-oriented CCR model to evaluate the efficiencies of the 12 supermarkets, we can illustrate the efficiency evaluation process on a 2-dimensional plane, as shown in Figure 1(a). Let DMU L be a unit that wants to improve, and let it be called the evaluated DMU. Then, the rest of the DMUs are called compared DMUs. DMUs A and B, which are located on the efficient frontier, belong to the reference set of DMU L, and HCU (Hypothetical Composite Unit) is suggested as benchmarking target of DMU L. The benchmarking target of DMU L is a hypothetical DMU that does not actually exist. This is the first deficiency of conventional DEA mentioned in Section 1. DMU A and DMU B, which are located on the efficient frontier, can be considered as benchmarking target of DMU L, but there are no reasonable ways to select benchmarking target between the two units for DMU L. This is the second deficiency of conventional DEA. DMU L has to reduce the 40 employees ($x1$) and 4000 m^2 floor area ($x2$) to benchmark HCU which is the benchmarking target for DMU L. However, it is very difficult for DMU L to reduce the floor area at once to benchmark the HCU. Likewise, it is somewhat unrealistic in that a very inefficient DMU cannot reach the benchmarking target at once, especially when the inefficient DMU is far from the benchmarking target. This is the third deficiency of conventional DEA. To resolve the third deficiency, the stratification method (Joe, 2003) illustrated in Figure 1(b) was proposed. If we apply this method directly to the supermarket example, we can obtain the five layers shown in the figure. DMU can improve its efficiency by crossing the sequence of layers represented by DMU K, H, F and B. Even though this solves the irrationality problem of conventional DEA, it does not consider various strategies for the selection of the benchmarking target.

Next subsection, in order to remedy the problem of not considering improving direction and unit similarity, an integrated benchmarking method combining direction and similarity together for the selection of the next step benchmarking target using DEA is discussed.



(a) Benchmarking target of conventional DEA (b) Benchmarking path by stratification method

Figure 1. Benchmarking target of a conventional DEA and a stratification method

3.1 Selection of benchmark targets considering direction and unit similarity

To find more effective benchmarking target, we consider two aspects: improving direction and unit similarity. First, improving direction is considered by selecting path with DMUs, which are located close to the line from the inefficient DMU to the final target. In Figure 2, the red line is indicated as L-1 presenting the direction of DMU L to benchmark DMU B, which is called a benchmarking line. If we consider direction to select next step benchmarking target for DMU L, DMU K, which is the closest to the benchmarking line, can be selected. In the present study, to determine closeness to the direction, we use the difference between two improvement ratios: improvement ratio for final benchmarking target (IRB) and improvement ratio for compared DMUs (IRC). The former is an improvement ratio of input factors for evaluated DMU to benchmark final benchmarking target, and the latter is an improvement ratio of input factors for evaluated DMU to benchmark compared DMUs.

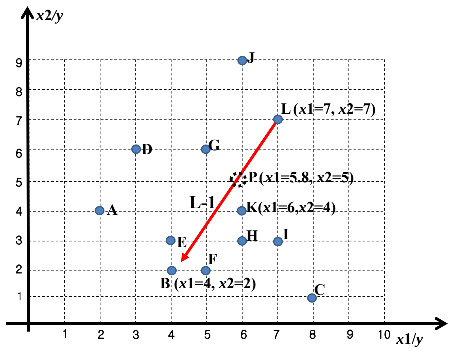


Figure 2. Direction of benchmarking from DMU L to DMU B

In order for DMU L to benchmark DMU B, input values ($x1$ and x2) of DMU L have to be reduced by 3 and 5, respectively, and the improvement ratio can be calculated as 0.6 (=3/5). If we select the DMU B as a final benchmarking target of DMU L, this ratio can be IRB. Let additional DMU P be a unit located on the benchmarking line. Then, the direction of DMU L to benchmark DMU P is exactly the same as that of the benchmarking line. The improvement ratio according to which DMU L can benchmark DMU P can be calculated as 0.6 (=1.2/2), and this ratio is IRD. This is equal to the improvement ratio necessary for DMU L to benchmark DMU B. With these results, it can be shown that two DMUs that have the same improvement ratio from the evaluated DMU share the same direction. So, if the improvement ratio of some DMUs is equal to the improvement ratio by which an evaluated DMU can benchmark a final benchmarking target, it can be determined that the improving directions are the same.

Based on this principle, a degree of direction can be defined as follows.

Definition 1: **Degree of direction** (δ_j)

A degree of direction of the j-th DMU (δ_j) can be explained as relative difference values of the two improvement ratios. The formulation for identification of δ_j can be represented by Equation (3).

$$\delta_j = 1 - \left[(\cos^{-1} \frac{\sum_{r=1}^{m} g_r h_r}{\sqrt{\sum_{r=1}^{m} g_r^2} \sqrt{\sum_{r=1}^{m} h_r^2}}) / \hat{\delta} \right] \tag{3}$$

where

g_r: Difference of the r-th input factors between evaluated DMU and final benchmarking target

h_r: Difference of the r-th input factors between evaluated DMU and compared DMU

$\hat{\delta}$: User-defined threshold value (Usually, it can be defined between 90 and 180)

δ_j indicates a relative inner product between g_r and h_r. DMU j, which has the maximum δ_j, means that it is closest to the direction from the evaluated DMU to the final benchmarking target.

Next, we consider unit similarity for selecting benchmarking path. Unit similarity is for the selection of similar DMUs to the evaluated DMU. Before we select a similar DMU, we classify all DMUs into similarity groups using SOM based on input patterns of the units. The closer distance between groups can be defined as the higher similarity. In order to calculate the similarity of each group, we set the distance value between DMUs in the same group be 0.5 and that for the adjacent group be 1.0. Based on this assumption, the distance can be calculated using Euclidean distance. We define degree of similarity to measure the similarity of the j-th DMU to the evaluated DMU in Definition 2.

***Definition 2*: Degree of similarity** (d_j)

A degree of similarity of the j-th DMU (d_j) is defined in Equation (4).

$$d_j = 1 - \frac{\sqrt{a^2 + b^2}}{P} \tag{4}$$

where

a: Horizontal distance between a group of an evaluated DMU and a group of compared DMUs.

b: Vertical distance between a group of an evaluated DMU and a group of compared DMUs.

P: Maximum distance among all groups (e.g. If the map generated by SOM is 4*4, P is 4.24(=$\sqrt{3^2+3^2}$).)

d_j in Equation (4) is the degree of similarity between a group the evaluated DMU belongs to and a group the compared DMUs belong to. DMU j with maximum value of d_j, is the most similar DMU to the evaluated DMU in terms of input patterns.

Improving direction and unit similarity above mentioned are used to select next benchmarking DMUs, when the evaluated DMU benchmarks the final benchmarking target. In order to select the final benchmarking target, we propose a new method, that is, it selects efficient DMU on the efficient frontier which is the closest to the ideal target. The ideal target is a point or a DMU where an evaluated DMU is expected to adopt in the production possibility set. The mathematical model of proposed method to select the final benchmarking target compatible with ideal target is represented by Equation (5).

$$\begin{aligned} & Max \sum_{r=1}^{s} w_r^y z_r - \sum_{i=1}^{m} w_i^x p_i \\ & s.t \quad \sum_{j=1}^{n} \lambda_{ij} x_{ij} = p_i k_i^x \quad i = 1,\dots,m \\ & \qquad \sum_{j=1}^{n} \lambda_{rj} y_{rj} = z_r k_r^y \quad r = 1,\dots,s \\ & \qquad k_i^x, k_r^y \in PPS, \ z_r \ge 1, \ p_i \le 1, \ w_r^y, w_r^x \ge 0 \end{aligned} \tag{5}$$

In Equation (5), k_i^x, k_r^y are the ideal target of i-th inputs and r-th outputs, w_i^x, w_r^y are the weights attached to the ideal target of i-th inputs and r-th outputs, and PPS means production possibility set. If optimal values for $p_{i,}$ z_r are p_i^*, z_r^*, the final benchmarking target of i-th inputs are $p_i^* k_i^x$ and r-th outputs are $z_r^* k_r^y$.

Evaluation method using combination of improving direction and unit similarity to select the next benchmarking DMU, a benchmarking evaluation function is described in Definition 3.

***Definition 3*: Benchmarking evaluation** (e_j)

A benchmarking evaluation function (e_j) to select benchmarking DMUs measures the possibility of being chosen as the next benchmarking target, and the formulation of the function is presented in Equation (6).

$$\begin{aligned} & e_j = \delta_j w_1 + d_j w_2, \quad j \in F(J) \\ & w_1, w_2 \ge 0, \ w_1, w_2 \le 1, \ w_1 + w_2 = 1 \end{aligned} \tag{6}$$

where

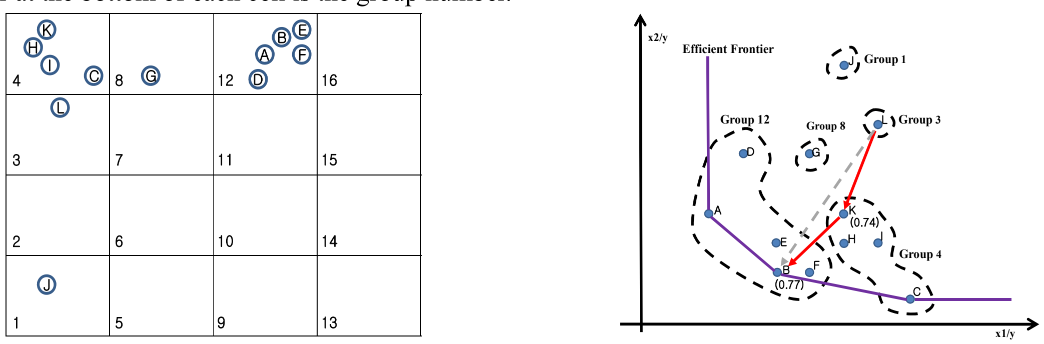
w_1,w_2: Weight values of the degree of direction and similarity.
J: A set of DMUs that is more efficient than a certain cut (α) from the evaluated DMU.
$F(J)$: Represents the correspondence from a DMU set to the corresponding subscript index set J.
α: User-defined minimum efficiency value to benchmark the next benchmarking DMU. ($0\leq \alpha \leq 1$)

w_1 and w_2 are weights assigned to the degree of direction and similarity, respectively. Different weights can be given to the degree of direction or similarity to impart more emphasis, and the sum of each weight must be equal to 1.

Based on the above definition and principles, the procedure of the integrated benchmarking method combining direction and similarity follows.

- Step 1: Measure relative efficiencies of all DMUs using DEA and select an evaluated DMU.
- Step 2: Select a final benchmarking target of the evaluated DMU.
- Step 3: Generate groups based on the similarity using SOM.
- Step 4: Evaluate δ_j, d_j and e_j.
- Step 5: Select a new benchmarking DMU as the next benchmarking target of the evaluated DMU by choosing the maximum value of e_j.
- Step 6: Substitute the new benchmarking target selected in Step 5 for the evaluated DMU.
- Step 7: If the new benchmarking target is equal to the final benchmarking target, terminate the procedure. Otherwise, go to Step 4.

For easy understanding of the benchmarking method discussed in this section, we apply the data from Table 1. After measuring the efficiencies of the DMUs, DMU L was determined to be an evaluated DMU. If DMU L is expected to reduce x1 and x2 by 2 and 4, the ideal target of DMU L can be determined as a point where x1 and x2 are 5 and 3, respectively. DMU B on the efficient frontier was selected as final benchmarking target of DMU L with the given values of weights w_1^x and w_2^x as 50 and 50, respectively. In order to generate similarity groups, we used SOM, of which the result is illustrated in Figure 3(a). The input variables of SOM are Floor Area and Employee of Table 1. Some condition variables of SOM are predetermined. The training cycle is set to 100, and the learning parameters for the start point and the end point are set to 0.9, and 0.1 respectively. The map configuration is set to 4 by 4. In Figure 3(a), the number at the bottom of each cell is the group number.



(a) Results of SOM from Table 1 (b) Benchmarking DMUs of proposed method

Figure 3. Results of SOM and proposed method

The minimum efficiency (α) was given a value of 0.1, and the w_1 and w_2 were 0.5 and 0.5, respectively. The results for the calculation of δ_j, d_j and e_j are illustrated in Table 2. The maximum value of e_j is DMU K in Table 2, and thus it is selected as the first benchmarking target of DMU L.

Table 2. Results for e_j from DMU L

DMU_j	A	B	C	D	E	F	G	H	I	K
δ_j	0.38	1.00	0.10	0.00	0.87	0.80	0.28	0.62	0.31	0.72
d_j	0.47	0.47	0.76	0.47	0.47	0.47	0.67	0.76	0.76	0.76
e_j	0.43	0.73	0.43	0.24	0.67	0.63	0.47	0.69	0.54	**0.74**

The evaluated DMU is substituted for DMU K, and similar procedure has been applied to the DMU K. The values of e_j among DMU A, B, E, F, which are more efficient than α from DMU K, are calculated. The results for the calculation of δ_j, d_j and e_j are illustrated in Table 3. Now, DMU B has the maximum value of e_j in Table 3, and it is determined as the second benchmarking target of DMU L. Since DMU B is equal to the final benchmarking target, the procedure can be terminated. Figure 3(b) show the result of stepwise benchmarking path of DMU L based on direction and similarity. That is, from the fact that DMU K exists in a group that is located closest to the group of DMU L, we can show that similarity concept is considered when we choose benchmarking target.

Table 3. Results for e_j from DMU K

DMU_j	A	B	E	F
δ_j	0.00	1.00	0.31	0.31
d_j	0.53	0.53	0.53	0.53
e_j	0.27	**0.77**	0.42	0.42

In applying our method to DMUs, resulting path can vary depending on SOM parameters, weight values assigned to direction and similarity, and minimum efficiency value. We summarized the effect of changing values of the parameters. First, as increasing (decreasing) values of the training cycle and learning parameters, the number of new benchmarking target is decreasing (increasing). Second, as increasing (decreasing) value of α, the number of benchmarking target is decreasing (increasing). Third, as increasing value of w_1 (w_2), there is a stress point on direction (similarity) to select next benchmarking target from evaluated DMU.

4. Case study

For case study, data have been collected for 21 East Asian container terminals, accessing relevant data sources from *Containerization International Year Book 2005*. We applied our method to the data set and we analyzed the result with different values of parameters such as α and weights assigned to the w_1 and w_2. First, the efficiencies of the container terminals were evaluated by simple DEA according to the numbers of berths, the lengths of berths (m), the total area of the port (km^2) and the number of cranes, as inputs, while the total container traffic (TEU) data were used as outputs. Four terminals (Hongkong, Sanghai, Shenzhen, Xiamen) were determined to be on the efficient frontier, and the remaining 18, inefficient. Next, to explain the selection of benchmarking path, we choose Kwangyang, as an evaluated DMU (efficiency score is 0.18), which is one of very inefficient DMUs. Shenzhen was selected as the final benchmarking target for Kwangyang port. We assigned the same value, 0.5, to both w_1 and w_2, and the value of 0 to α, respectively. From the results for the calculation of δ_j, d_j and e_j from Kwangyang, Tianjin, which has the maximum value of e_j, 0.87, among the compared DMUs, is selected as the first benchmarking target of Kwangyang. Then, Tianjin is substituted as the evaluated DMU. From the results for the calculation of δ_j, d_j and e_j from Tianjin, Lianyungang is selected as the second benchmarking target of Kwangyang. Then, Lianyungang is substituted as the evaluated DMU. Shenzhen, which has the maximum value of e_j, 0.91, is selected as the third benchmarking target of Kwangyang. Since Shenzhen is the same as the final benchmarking target of Kwangyang, the procedure is terminated. The results for the calculation of δ_j, d_j and e_j are shown in Table 4.

Table 4. Results for e_j from Kwangyang

DMU	Hong kong	Sang hai	Shen zhen	Bu san	Kaoh siung	Qing dao	Ning bo	Tian jin	Guan gzhou	To kyo	Xiam en	Da lian	Nag oya	Osa ka	Keel ung	Inch eon	Lianyun gang
δ_j	0.93	1.00	0.87	0.92	0.69	0.93	0.98	0.69	0.93	0.68	0.90	0.48	0.43	0.24	0.22	0.60	0.74
d_j	0.75	0.61	0.61	0.61	0.61	0.61	0.75	0.61	0.75	0.61	0.61	0.82	0.61	0.91	0.75	0.82	0.75
e_j	0.73	0.84	0.80	0.74	0.76	0.65	0.77	**0.87**	0.65	0.64	0.75	0.65	0.52	0.58	0.49	0.71	0.75
Result for e_j from Tianjin																	
δ_j	0.99	0.00	1.00	0.00	0.99	0.95	1.00	-	-	-	-	-	-	-	0.93	-	0.98
d_j	0.82	0.91	0.82	0.47	0.82	0.47	0.82	-	-	-	-	-	-	-	0.65	-	0.91
e_j	0.91	0.46	0.91	0.23	0.91	0.71	0.91	-	-	-	-	-	-	-	0.79	-	**0.94**
Result for e_j from Lianyungang																	
δ_j	0.99	0.00	1.00	-	0.91		0.96	-	-	-	-	-	-	-	0.40	-	-
d_j	0.82	0.91	0.82	-	0.82		0.82	-	-	-	-	-	-	-	0.65	-	-
e_j	0.90	0.46	**0.91**	-	0.87		0.89	-	-	-	-	-	-	-	0.52	-	-

If we assign very small value to α, Tianjin and Lianyungang can be selected as the first and second benchmarking targets of Kwangyang. However, the benchmarking may not be useful since the efficiency gap of Tianjin and Lianyungang is 0.08, and it's too small (Efficiencies of Tianjin and Lianyungang are 0.61 and 0.69, respectively). In this case we can set larger value of α to avoid such unrealistic selection. The results of stepwise benchmarking path from Kwangyang according to different values of α are shown in Table 5.

Table 5. Results of the new benchmarking path of Kwangyang according to different values of α

α	Benchmarking path of Kwangyang	α	Benchmarking path of Kwangyang
0	Kwangyang → Tianjin → Lianyungang → Shenzhen	0.5	Kwangyang → Tianjin → Shenzhen
0.1	Kwangyang → Tianjin → Shenzhen	0.6	Kwangyang → Shenzhen
0.2	Kwangyang → Tianjin → Shenzhen	0.7	Kwangyang → Shenzhen
0.3	Kwangyang → Tianjin → Shenzhen	0.8	Kwangyang → Shenzhen
0.4	Kwangyang → Tianjin → Shenzhen	0.9	Kwangyang

Next, we found benchmarking path with different pairs of weights for direction and similarity. The result of finding new benchmarking targets for Kwangyang can be different according to the values assigned to the w_1 and w_2. When we assigned the value, 0.2, to α, the results of the benchmarking path of Kwangyang according to the change of the values assigned to the w_1 and w_2 are shown in Table 6.

Table 6. Results of the benchmarking path of Kwangyang according to the change of the w_1 and w_2

w_1	w_2	Stepwise benchmarking target	w_1	w_2	Stepwise benchmarking target
0.1	0.9	Kwangyang → Dalian → Tianjin → Shenzhen	0.6	0.4	Kwangyang → Tianjin → Shenzhen
0.2	0.8	Kwangyang → Tianjin → Shenzhen	0.7	0.3	Kwangyang → Tianjin → Shenzhen
0.3	0.7	Kwangyang → Tianjin → Shenzhen	0.8	0.2	Kwangyang → Tianjin → Shenzhen
0.4	0.6	Kwangyang → Tianjin → Shenzhen	0.9	0.1	Kwangyang → Shenzhen
0.5	0.5	Kwangyang → Tianjin → Shenzhen	1.0	0.0	Kwangyang → Shenzhen

5. CONCLUSIONS

In this paper, we proposed a DEA based method of selecting stepwise benchmarking path, which takes both direction and similarity into account. This is a new method that is formulated to remedy the drawback of the conventional DEA and the existing stepwise benchmarking, which is, that they consider only the influence of efficiency when an inefficient DMU has to benchmark its benchmarking target using DEA. Our approach differs from previous ones in that it considers two strategies of direction and similarity together in rendering inefficient DMUs efficient. For direction, we defined a target selection method that finds the closest efficient DMU in the direction in which the inefficient DMU to benchmark the final benchmarking target. For similarity, we defined a target selection method that finds the similar DMUs with the evaluated DMU based on the input patterns using SOM. As an application of the proposed method, benchmarking of an East Asian container terminal was tested in the present study. The results show that the stepwise benchmarking target of an inefficient DMU could be found. This method is effective also in that it can suggest alternative benchmarking paths and targets by changing the condition variables, the value of minimum efficiency and weights used for direction and similarity.

ACKNOWLEDGMENT
This research is supported by Center for International Exchange & Cooperation on Shipping, Port and International Logistics founded by Ministry of Land, Transport and Maritime Affairs.

6. REFERENCES

Alirezaee, M. R. and Afsharian, M. (2007). Model improvement for computational difficulties of DEA technique in the presence of special DMUs. *Applied mathematics and Computation*, 186: 1600-1611.
Ammons, D. N. (2002), Benchmarking as a performance management tool: experiences among municipalities in North Carolina. *European Journal of Operational Research*, 140: 249-265.
Bogetoft, P., & Hougaard, J. L. (1999). Efficiency evaluations based on potential improvements. *Journal of Productivity Analysis*, 12: 233–247.
Charnes, A., Cooper, W. W. and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2: 429-444.
Cooper, Willian W., Lawrence M. Seiford and Kaoru Tone (2006). Introduction to Data Envelopment Analysis and Its uses: with DEA solver software and reference. *Interface.*
Donthu, N., Hershberger, E. K. and Osmonbekov, T. (2005). Benchmarking marketing productivity using data envelopment analysis. *Journal of Business Research*, 58: 1474-1482.

Gonzales, E., & Alvarez, A. (2001). From efficiency measurement to efficiency improvement: The choice of a relevant benchmark, *European Journal of Operational Research*, 133: 512–520.
Grupp, H. (1990). Technometrics as a missing link in science and technology indicator. *Measuring the Dynamics of Technological Change*, 57-76.
Haykin S. (1999). Neural networks a comprehensive foundation. *Prentice Hall*, 1: 443-483.
Joe Zhu (2003). *Quantitative models for performance evaluation and benchmarking-Data Envelopment Analysis with Spreadsheets and DEA Excel Solver*. Kluwer Academi Publishers.
Kohonen. T. (1988). An introduction to neural computing. *Neural Networks*, 1: 3-16.
Park, J-H., Bae, H-R. and Lim, S-M. (2010). Method of benchmarking route choice based on the input-similarity using DEA and SOM. *Journal of the Korean Institute of Industrial Engineers*, 36: 32-41.
Ross, A., Droge, C. (2002). An integrated benchmarking approach to distribution center performance using DEA modeling. *Journal of Operations Management*, 20: 19-32.
Tata, J., Prasad, S. and Motwani, J. (2000). Benchmarking quality management practices: U.S. Versus Costa Rica. *Multinational Business Review*, 8: 37-51.
Shaneth A. E., Hee, S-S., Young, A-K., Su, H-N. and Shin, C-K. (2009). A method of stepwise benchmarking for inefficient DMUs based on the proximity-based target selection. *Expert Systems with Applications,* 36: 11595-11604.
Thanassoulis, E., & Dyson, R. G. (1989). Estimating preferred target input–output levels using data envelopment analysis. *European Journal of perational Research*, 56: 80–97.

Session B4: Shipping

·Day1: Sep. 15 (Wed.)

·Time: 13:20 - 14:40

·Chair: Hwa-Joong Kim

·Room: Grand Ballroom, 5F

LINER SHIPPING FLEET DEPLOYMENT WITH CARGO TRANSSHIPMENT AND DEMAND UNCERTAINTY

Shuaian Wang[1] and Qiang Meng[2*]

[1]Department of Civil Engineering
National University of Singapore
Singapore, 117576
e-mail: swang@nus.edu.sg

[2] Department of Civil Engineering
National University of Singapore
Singapore, 117576
e-mail: cvemq@nus.edu.sg

Abstract: This paper addresses a novel liner shipping fleet deployment problem characterized by cargo transshipment, multiple container routing options and uncertain demand, with the objective of maximizing the expected profit. This problem is formulated as a stochastic program and solved by the sample average approximation method. In this technique the objective function of the stochastic program is approximated by a sample average estimate derived from a random sample, and then the resulting deterministic program is solved. This process is repeated with different samples to obtain a good candidate solution along with the statistical estimate of its optimality gap. We apply the proposed model to a case study inspired from real-world problems faced by a major liner shipping company. Results show that the case is efficiently solved to 1% of relative optimality gap at 95% confidence level.

Key Words: Liner shipping; Fleet deployment; Transshipment; Demand uncertainty; Sample average approximation

1. INTRODUCTION

Fleet deployment decisions arise at the tactical planning level of liner shipping networks. Fundamentally, the liner shipping fleet deployment problem (LSFDP) involves the allocation of ships to service routes for delivering containers. To maximize the total profit, liner shipping companies need to assess the trade-off between the cost and capacity of ships. The LSFDP has three complicating factors due to the unique characteristics of liner shipping services: (i) a certain service frequency or a minimum level of service frequency must be maintained; (ii) more than one routing option is available for delivering containers from origin to destination; (iii) fleet deployment decisions have to be made prior to knowing the exact shipping demand.

Various pure or mixed-integer linear programming models for the LSFDP have been developed to account for different levels of practical constraints arising in liner shipping operations. Perakis and Jaramillo (1991) and Jaramillo and Perakis (1991) contributed the pioneering work for the LSFDP. They proposed a linear programming formulation incorporating the capacity constraint, minimum service frequency requirement and ship chartering. The cargo amount between pairs of ports is given and the aim is to minimize the total fleet operating cost including fuel costs, daily running costs, port charges and canal fees. The aforementioned linear programming model assumes that the number of ships of each type allocated to a route is a continuous rather than an integer variable. To remedy this unrealistic assumption, Powell and Perakis (1997) presented an integer linear programming model for the LSFDP. The above mentioned studies all assume that cargo demand between pairs of ports on each service route is *a priori* known, Cho and Perakis (1996) relaxed this assumption by formulating a model in which the given cargo demand between two specific ports can be served by any route passing through these two ports. Nevertheless, cargo transshipment is still not allowed in this model. To take into account cargo transshipment, Mourão et al. (2001) studied the LSFDP in a simple hub-and-spoke (H&S) network which consists of two routes (a feeder route and a main route) and one origin-destination (O-D) port pair. All containers must be transshipped at the hub port in the feeder route. As the cruising speed of ships have direct implications on the round-trip time and bunker consumption, Gelareh and Meng (2010) presented a mixed-integer programming model by discretizing the cruising speed to integrate the decision of optimal ship speed into the LSFDP. Unlike the above deterministic models, in a recent contribution, Meng and Wang (2010) investigated the uncertain demand in the LSFDP. They assume that a certain level of service for each route has to be maintained. The level of service is defined as the probability that all cargo demand on the route can be satisfied. Thus the uncertain demand is formulated as chance constraints and these chance constraints are then transformed into their equivalent deterministic constraints.

The LSFDP models found in the literature have two main shortcomings when compared to the practical liner shipping operations. First, most of the models assume that the cargo demand is associated with a certain liner route, thus can solely be fulfilled by direct transportation services on the route without transshipment. As a matter of fact, in liner shipping practices, it is prevalent to transfer containers at an intermediate port from one ship to another. Transshipment not only enables the use of large containerships to exploit the economies of scale in the hub-and-spoke networks, but also expands the service scope in that the number of O-D pairs can be much more than the number of direct O-D services operated. Moreover, many routing options for an O-D pair become available due to the possibility of transshipment. Mourão et al. (2001) considered the issue of transshipment in a very special network setting with only one routing option. Cho and Perakis (1996) investigated multiple routing options without transshipment. These two models cannot be adapted to accommodate general cases of both transshipment and multiple routing options. Second, few attempts have been devoted to the inclusion of uncertain demand in the fleet deployment decisions. In contrast to industrial and tramp shipping which operate schedules in response to actual demand, liner shipping is based on a fixed schedule which is generally published up to six months into the future. This means that the fleet deployment decisions are made depending on the demand forecast. The forecast is seldom error-free and hence it is advisable to capture the uncertain nature of the demand. As will be reported in Section 5, compared with simply using the average demand value, the inclusion of the uncertain demand in making the fleet deployment decisions yields superior solutions in terms of the expect profit. The only model with uncertain demand we are aware of is proposed by Meng and Wang (2010). However, this model relies on the assumption that the container numbers in the O-D port pairs are independent and normally distributed. Thus its applicability to real-life situations is rather limited.

In this paper, we take the initiative to investigate the LSFDP with cargo transshipment, multiple routing options and demand uncertainty. The uncertain demand is represented by a multivariate random variable and can have any general probability distribution. This problem is formulated as a stochastic model with the objective of maximizing the expected profit. We use the sample average approximation (SAA) method to solve the model and estimate the statistical optimality gap.

The remainder of this paper is organized as follows: Section 2 is the detailed description of the LSFDP. In Section 3 we first present the stochastic programming formulation for the basic LSFDP; then we discuss some extensions of the model to adapt to more practical constraints. Section 4 is dedicated to employing the SAA method to address the stochastic model. Section 5 is a numerical example to assess the proposed model. Conclusions are presented in Section 6.

2. PROBLEM DESCRIPTION

Consider a liner shipping company operating a set of routes to transport containers between pairs of ports. The itinerary (sequence of portcalls) of each route is given as the input for the fleet deployment problem. All the routes have a weekly service frequency. The liner shipping company deploys a string of ships of the same type on each route in order to maintain the weekly service frequency and to deliver containers at maximum profit. Liner routes intersect at the common port of call, and thus containers can be transshipped between ships on two liner routes. An illustrative liner shipping network is shown in Figure 1.

In the mathematical description of the problem, the set of routes is represented by R. The number of portcalls in a round-trip of a route $r \in R$ is denoted by N_r. We use the port calling sequence on a route to refer to the port of call because a port may be visited more than once during a round-trip. For example, route r_3 in Figure 1 has 4 portcalls. If SH is defined as the 1st portcall, both the 2nd and the 4th portcalls refer to SG, but they are different portcalls. Similarly, we define the i^{th} leg of a route r, denoted by $r(i)$, as the voyage from the i^{th} portcall to the $(i+1)^{th}$, except that the N_r^{th} leg is from the N_r^{th} portcall to the 1st . We further define $I_r:=\{1,2 \ldots N_r\}$. A corresponding definition for the routes in Figure 1 is presented in Table 1.

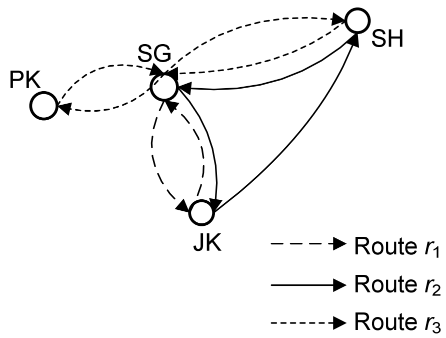


Figure 1. A Liner Shipping Network
SH: Shanghai; SG: Singapore; JK: Jakarta; PK: Port Klang

Table 1. Portcalls and Legs of Routes

Route	Portcalls	Legs
r_1: JK-SG	1st: JK; 2nd: SG	$r_1(1)$: JK-SG; $r_1(2)$: SG-JK
r_2: SH-SG-JK	1st: SH; 2nd: SG; 3rd:JK	$r_2(1)$: SH-SG; $r_2(2)$: SG-JK; $r_2(3)$: JK-SH
r_3: SH-SG-PK-SG	1st: SH; 2nd: SG; 3rd:PK; 4th: SG	$r_3(1)$: SH-SG; $r_3(2)$: SG-PK; $r_3(3)$: PK-SG; $r_3(4)$: SG-SH

2.1 Fleet Deployment

The liner shipping company has a set of ships to deploy on the liner routes. Let S be the set of ship types and the container capacity in terms of twenty-foot-equivalent unit (TEU) of a particular ship type $s \in S$ is denoted by V_s. We assume that there are an unlimited number of ships in each type. When a route r is deployed with ships of type s, we first assume that the number of ships is fixed and we will relax this assumption later on. Consequently, the cost of operating the route, denoted by c_{rs} ($/week), including the fuel costs, daily running costs, port charges and canal fees as in Perakis and Jaramillo (1991), is also determined. As each route has a weekly service frequency, the round-trip time equals the number of ships deployed multiplied by 7 days. For instance, if the round-trip time is 35 days on a route r, 5 ships are required to ensure the same portcall is visited on the same day every week. The round-trip time consists of the time at sea (the round-trip distance divided by cruising speed), the time for pilotage in and out of ports, and the berth time for container handling. We assume the ship speed and the pilotage time at port are given, thus the maximum berth time T_{rs} (h) for container handling at all portcalls of the route-ship type combination r-s is also determined and must be respected.

2.2 Container Routing Plan

The liner shipping company transport containers from the origin port to the destination. Denote the set of O-D port pairs by D and lowercase letter d refers to a particular O-D. The number of containers for an O-D d is denoted by a random variable ξ_d (TEUs/week). We assume that the value of ξ_d, although unknown in advance, remains constant in each week over the fleet deployment planning period. The joint probability distribution of ξ_d, $d \in D$, is estimated e.g. from historical data and used as input for the fleet deployment decisions. The revenue from delivering one TEU for the O-D d is g_d (dollars/TEU).

Usually there are many routing options for delivering containers from origin to destination. For example, in Figure 1, containers from JK to SH can be transported (i) on route r_2; or (ii) on route r_1 to SG, and then transshipped to ships on route r_3 and then delivered to SH. Each routing option is defined as a *container routing plan* (referred to as *routing plan* for short hereafter). A routing plan, denoted by h, consists of a group of ordered legs to be visited. For example, the above mentioned two routing plans for the demand JK to SH, can be denoted by $r_2(3)$ and $r_1(1) \rightarrow r_3(4)$, respectively. Some routing plans for the liner shipping network in Figure 1 are provided in Table 2.

Table 2. Routing Plans for Different O-Ds

O-D	Routing Plans
d_1: JK-SH	h_1: $r_2(3)$
	h_2: $r_1(1)\rightarrow r_3(4)$
d_2: SH-PK	h_3: $r_3(1)\rightarrow r_3(2)$
d_3: SH-SG	h_4: $r_2(1)$
	h_5: $r_3(1)$

A routing plan contains all information on how containers are transported, such as origin, destination, liner routes involved, and transshipment ports. An O-D d may have several routing plans, the set of which denoted by H_d, whereas a routing plan corresponds to a single O-D. Containers of the same O-D can be split among several routing plans. The set of all the routing plans, denoted by H, is provided by the liner shipping company as an input for the fleet deployment model.

The operating cost of a route is determined once its fleet deployment decision is made. The main variable cost in liner shipping operations is the container handling cost, which consists of loading cost at origins, discharging cost at destinations and transshipment costs at transshipment ports, if any. Two routing plans for the same O-D may incur different handling costs. For example, the routing plan h_1 in Table 2 involves exclusively the loading cost at JK and discharging cost at SH, while the routing plan h_2 is associated with an additional transshipment cost at SG. We denote the handling cost of the routing plan h by c_h (dollars/TEU).

Routing plans also influence the berth time of related routes. Let t_{ps} (h/TEU) be the average container handling time at port p for ship type s. Note that larger ships allow more quay cranes to work simultaneously, thus having a higher handling efficiency. Containers in routing plan h_2 are loaded at JK and discharged for transshipment at SG on route r_1. Hence if route r_1 is deployed with ships of type s, let p=JK and q=SG, then the *incremental* berth time on route r_1 for routing plan h_2 is $t_{ps}+t_{qs}$ (h/TEU).

3. MATHEMATICAL MODEL

In view of the uncertain demand, the liner shipping company aims to maximize its expected profit. We first formulate the expected value model for the basic LSFDP described in Section 2. Then we discuss some extensions of the model to handle more practical constraints.

3.1 Stochastic Model for the Basic LSFDP

To formulate the stochastic model of the basic LSFDP, we need further notations. Let the binary coefficient δ_{hri} indicate whether routing plan h contains the i^{th} leg of route r (δ_{hri}=1) or not (δ_{hri}=0). Denote by t_{hrs} (h/TEU) the incremental berth time incurred by one TEU in routing plan h on route r deployed with ships of type s. t_{hrs}=0 if the routing plan h does not use any leg of the route r. The decision variables are as follows: f_h is the number of TEUs transported in routing plan h; x_{rs} is a binary variable which takes the value 1 if and only if ships of type s are deployed on route r. The expected value model for the basic LSFDP is:

[EVM]
$$z^* = \max \mathbb{E}\left[\sum_{d\in D}\sum_{h\in H_d}\left(g_d - c_h\right)f_h - \sum_{r\in R}\sum_{s\in S} c_{rs}x_{rs}\right] \quad (1)$$

s.t.
$$\sum_{h\in H_d} f_h \le \xi_d, \forall d \in D \quad (2)$$

$$\sum_{h\in H}\delta_{hri}f_h - \sum_{s\in S}V_s x_{rs} \le 0, \forall r \in R, \forall i \in I_r \quad (3)$$

$$\sum_{h\in H} f_h \sum_{s\in S} t_{hrs}x_{rs} \le T_{rs}x_{rs}, \forall r \in R \quad (4)$$

$$\sum_{s\in S} x_{rs} = 1, \forall r \in R \quad (5)$$

$$x_{rs} \in \{0,1\}, \forall r \in R, \forall s \in S \tag{6}$$

$$f_h \geq 0, \forall h \in H \tag{7}$$

The objective function (1) maximizes the expected profit. The first term is the revenue less the handling cost from shipping containers; the second term is the operating cost of the routes. Constraints (2) require that the number of containers transported do not exceed the demand. Constraints (3) are the capacity constraints imposed on each leg of the routes. Constraints (4) ensure that the maximum berth time is respected. Constraints (5) enforce that exactly one type of ship is deployed on each route. Constraints (6) denote x_{rs} as binary variables and constraints (7) denote f_h as nonnegative continuous variables.

3.2 Model Extensions

EVM has some simplifying assumption and now we extend **EVM** so that it can accommodate more practical constraints.

We have assumed that there are an unlimited number of ships in each type. If this assumption is violated, let n_s be the maximum number of ships of type s available and n_{rs} be the number of ships required if ships of type s are deployed on route r, then we can add the following constraints:

$$\sum_{r \in R} n_{rs} x_{rs} \leq n_s, \forall s \in S \tag{8}$$

It is very probable that the liner shipping company not only services the spot market where the cargo demand is optional, but also has some contracted cargo which must be transported. Denote by m_d (TEUs) the number of contracted cargo for O-D d, the following constraints hold:

$$\sum_{h \in H_d} f_h \geq m_d, \forall d \in D \tag{9}$$

Practical liner routes normally provide weekly services. Otherwise we may let t_r be the service frequency (number of days between two arrivals at a portcall) of route r, and constraints (3) and (4) can be restated as:

$$\frac{t_r}{7} \sum_{h \in H} \delta_{hri} f_h - \sum_{s \in S} V_s x_{rs} \leq 0, \forall r \in R, \forall i \in I_r \tag{10}$$

$$\frac{t_r}{7} \sum_{h \in H} f_h \sum_{s \in S} t_{hrs} x_{rs} \leq T_{rs} x_{rs}, \forall r \in R \tag{11}$$

Some types of ship might not be compatible with all the routes, e.g., due to physical restrictions at ports. For any r-s of the incompatible route-ship combinations, we simply define that $x_{rs}=0$.

When a liner route r is deployed with ships of type s, we have assumed that the number of ships is determined. If the liner shipping company decides to add one more ship in order to provide more berth time (here one week for a weekly service) for container handling at ports, we can simply consider this deployment as having $n_{rs}+1$ ships of type s_1 which have exactly the same properties as ships of type s.

In constraint (3), we have used the vessel capacity V_s as the leg capacity for all voyage legs. In practice, ships may not be allowed to fully load on some legs because of restricted water. Therefore the capacity might vary on different legs of a route. We can let V_{rsi} be the capacity on the i^{th} leg of route r when it is deployed with ships of type s, and constraints (3) can be restated as:

$$\sum_{h \in H} \delta_{hri} f_h - \sum_{s \in S} V_{rsi} x_{rs} \leq 0, \forall r \in R, \forall i \in I_r \tag{12}$$

4. SOLUTION ALGORITHM

In this section we present the solution algorithm for the basic model of the LSFDP. The extensions of the basic model can be addressed similarly. The basic model is difficult in that it is a stochastic one with nonlinear constraints (4). In

Section 4.1, we use the sample average approximation method to transform the stochastic model into an approximating deterministic model. Section 4.2 is dedicated to the statistical analysis of the solution quality. An equivalent linear formulation of constraints (4) is presented in Section 4.3.

4.1 Sample Average Approximation

Unless the uncertain demand ξ_d, $d\in D$, has a small number of possible realizations (scenarios), it is usually impossible to solve **EVM** exactly. One approach for approximately solving **EVM** is the sample average approximation (SAA) method (Verweij et al., 2003). The SAA method is an approach for solving stochastic optimization problems by using Monto Carlo simulation. In this technique the objective function of the stochastic program is approximated by a sample average estimate derived from a random sample. The resulting sample average approximating problem is then solved by deterministic optimization approaches. This process is repeated with different samples to obtain a good candidate solution along with the statistical estimate of its optimality gap.

To solve **EVM**, we first uses the Monte Carlo procedure to generate N independent and identically distributed observations ξ_d^n, n=1, 2… N, according to the joint probability distribution of the demand. Then the expected value function is approximated by the sample average function. Let I_N:={1, 2… N}, and the approximating deterministic model is:

[ADM]
$$z_N = \max \frac{1}{N}\sum_{n\in I_N}\sum_{d\in D}\sum_{h\in H_d}(g_d - c_h)f_h^n - \sum_{r\in R}\sum_{s\in S} c_{rs}x_{rs} \tag{13}$$

s.t.
$$\sum_{h\in H_d} f_h^n \le \xi_d^n, \forall d\in D, \forall n\in I_N \tag{14}$$

$$\sum_{h\in H}\delta_{hri}f_h^n - \sum_{s\in S}V_s x_{rs} \le 0, \forall r\in R, \forall i\in I_r, \forall n\in I_N \tag{15}$$

$$\sum_{h\in H} f_h^n \sum_{s\in S} t_{hrs}x_{rs} \le T_{rs}x_{rs}, \forall r\in R, \forall n\in I_N \tag{16}$$

$$\sum_{s\in S} x_{rs} = 1, \forall r\in R \tag{17}$$

$$x_{rs}\in\{0,1\}, \forall r\in R, \forall s\in S \tag{18}$$

$$f_h^n \ge 0, \forall h\in H, \forall n\in I_N \tag{19}$$

4.2 Statistical Analysis of Solution Quality

ADM is a deterministic optimization problem and its optimal objective value z_N and optimal fleet deployment decisions, denoted by the vector $\mathbf{x}_N$, can be used as an estimate of their counterparts z^* and $\mathbf{x}^*$ in **EVM**. This approach is justified by the epi-convergence theory: under mild assumptions, $\{z_N\}_{N=1}^{\infty}$ converges to z^* with probability 1, and accumulation points $\{\mathbf{x}_N\}_{N=1}^{\infty}$ are optimal solutions to **EVM** with probability 1 (Shapiro, 1991).

As **ADM** is an approximating model, we need to assess the quality of the solution provided by **ADM**. To this end, a common way is to estimate the lower and upper bounds of z^* in **EVM**. We use the method proposed by Mak et al. (1999) to estimate the statistical upper and lower bounds of z^*. To keep this paper relatively self-contained, we briefly introduce this method.

From Mak et al. (1999), the expected value of z_N has the property $\mathbb{E}z_N \ge \mathbb{E}z_{N+1} \ge z^*$ in the maximization problem **ADM**. Consequently, if we generate M independent samples of the uncertain demand, each of size N, and obtain M optimal objective values z_N^k and the optimal solution $\mathbf{x}_N^k$ for **ADM**, $k=1,2\cdots M$, then a statistical upper bound for z^* can be estimated by $\overline{U} = \sum_{k=1}^{M} z_N^k / M$. Let S_U be the sample variance of z_N^k, $k=1,2\cdots M$, then the statistic $\varepsilon_U = (\overline{U} - \mathbb{E}z_N)/\sqrt{S_U/M}$ has a t-distribution with $M-1$ degrees of freedom.

Let $t_{\alpha,M-1}$ satisfy $P\{T \le t_{\alpha,M-1}\} = 1-\alpha$, where the random variable T has a t-distribution with $M-1$ degrees of freedom, and $0 \le \alpha \le 1$, then $P\{Ez_N \le \overline{U} + t_{\alpha,M-1}\sqrt{S_U / M}\} = 1-\alpha$ and hence $\overline{U} + t_{\alpha,M-1}\sqrt{S_U / M}$ is an upper bound of z^* with at least $1-\alpha$ level of confidence.

To estimate the lower bound, let $k^* := \arg\max z_N^k$, $k = 1,2\cdots M$, and define $\mathbf{x}_N^* := \mathbf{x}_N^{k^*}$. Denote by z' the optimal objective value of **EVM** when the values of x_{rs} are fixed at $\mathbf{x}_N^*$. As $\mathbf{x}_N^*$ is a feasible solution to **EVM**, z' is a lower bound of z^*. Generate another N' ($N' \gg N$) observations $\boldsymbol{\xi}^j$, $j = 1,2\cdots N'$, according to the joint probability distribution of the demand. These N' observations are independent from the previous ones generated for the estimation of the statistical upper bound. Let z'_j be the optimal objective value of **EVM** when the values of x_{rs} are fixed at $\mathbf{x}_N^*$, and the random variables ξ_d are replaced by the observation $\boldsymbol{\xi}^j$, $j = 1,2\cdots N'$. It should be noticed that now **EVM** has become a deterministic optimization problem with simply the container routing variables f_h. Consequently $\overline{L} = \sum_{j=1}^{N'} z'_j / N'$ can be used to estimate z'. Similarly, denote by S_L the sample variance of z'_j, $j = 1,2\cdots N'$, then the statistic $\varepsilon_L = (\overline{L} - z') / \sqrt{S_L / N'}$ has a t-distribution with $N'-1$ degrees of freedom. Hence $\overline{L} - t_{\alpha,N'-1}\sqrt{S_L / N'}$ is a lower bound of z^* with at least $1-\alpha$ level of confidence.

Hence the optimal objective value of **EVM** z^* lies in the interval $[\overline{L}, \overline{U}]$ with at least $1-2\alpha$ level of confidence. To assess the quality of the candidate solution $\mathbf{x}_N^*$, we can use the relative optimality gap which is defined as $|(\overline{U} - \overline{L}) / \overline{L}|$.

4.3 Linearization

There are still difficulties in solving **ADM** as the constraints (4) are nonlinear. However, they can be replaced with the following equivalent linear constraints:

$$\sum_{h \in H} f_h t_{hrs} \le T_{rs} + M_H (1 - x_{rs}), \forall r \in R, \forall s \in S \tag{20}$$

In constraint (20), M_H is a huge number which ensures that when $x_{rs} = 0$, no constraints on the f_h variables are imposed. M_H can be set as the maximum container handling time during a round-trip of any route with any type of ship. To represent the value of M_H mathematically, let t_{rsi} denote the average container handling efficiency (h/TEU) at the i^{th} port of route r for ships of type s, then we can set $M_H = \max\{2V_s \sum_{i \in I_r} t_{rsi} \mid \forall r \in R, \forall s \in S\}$, in which $2V_s \sum_{i \in I_r} t_{rsi}$ means the berth time of route r with ships of type s, when full shipload of containers are discharged and another full shipload loaded at all portcalls.

5. CASE STUDY

We test our model and algorithm on a case inspired from real-world problems faced by a major liner shipping company. This test case has a network of 46 ports as shown in Figure 2. There are a total of 3 types of ship, 1500 TEUs, 3000 TEUs and 5000 TEUs, to be deployed on 11 routes that connect the 46 ports. A total of 652 O-D port pairs have non-zero container shipping demand. For each O-D d, we assume that there are three possible scenarios of the container number, μ_d, $\mu_d(1+\kappa)$, and $\mu_d(1-\kappa)$, $0<\kappa<1$, each having a probability of 1/3. The container numbers of different O-D pairs are independent. We test three cases of demand variability with $\kappa=0.2$, $\kappa=0.5$ and $\kappa=0.9$. In each case, the total number of possible realizations of the uncertain demand is 3^{652}.

Recall from Section 4.2 that the SAA method requires the solution of **ADM** M times, each including N sampled scenarios. In this case study we set M=21 and N=30. The number of observations N' for calculating the lower bound is set to be 201. We use the open source code lp_solve version 5.5 (Lp_solve, 2010) and develop the algorithm with C++ on Visual Studio 2005. All computations are carried out on a computer with an Intel Duo 3.20 GHz processor and 3 GB of RAM.

Computational results show that it takes less than 9 minutes to calculate one instance of **ADM**. The statistical lower bound, upper bound, confidence interval and relative optimality gap for different demand variability are presented in Table 3. In all the three cases of demand variability, the standard deviations of z'_j and z_N^k are less than 4% of their respective average value. Consequently the 95% confidence interval is very tight, and the relative optimality gap is less than 1%. Moreover, the variance of the statistical lower bound, upper bound, the width of the confidence interval and the relative optimality gap increase with the variability of the demand.

To justify the efforts for solving the stochastic fleet deployment model, we also calculate the expected profit of fleet deployment decisions without considering demand uncertainty with the following procedure: firstly, solve the deterministic optimization problem **EVM** by replacing the random variable ξ_d with their mean value μ_d in constraints (2), and get the optimal fleet deployment result; then use different realization of ξ_d to calculate the expected value of the profit similar to the calculation of lower bound in Section 4.2. The comparison result under different demand variability is presented in Figure 3. In all the demand variability cases, the expected profit with the inclusion of demand uncertainty in fleet deployment decisions is higher.

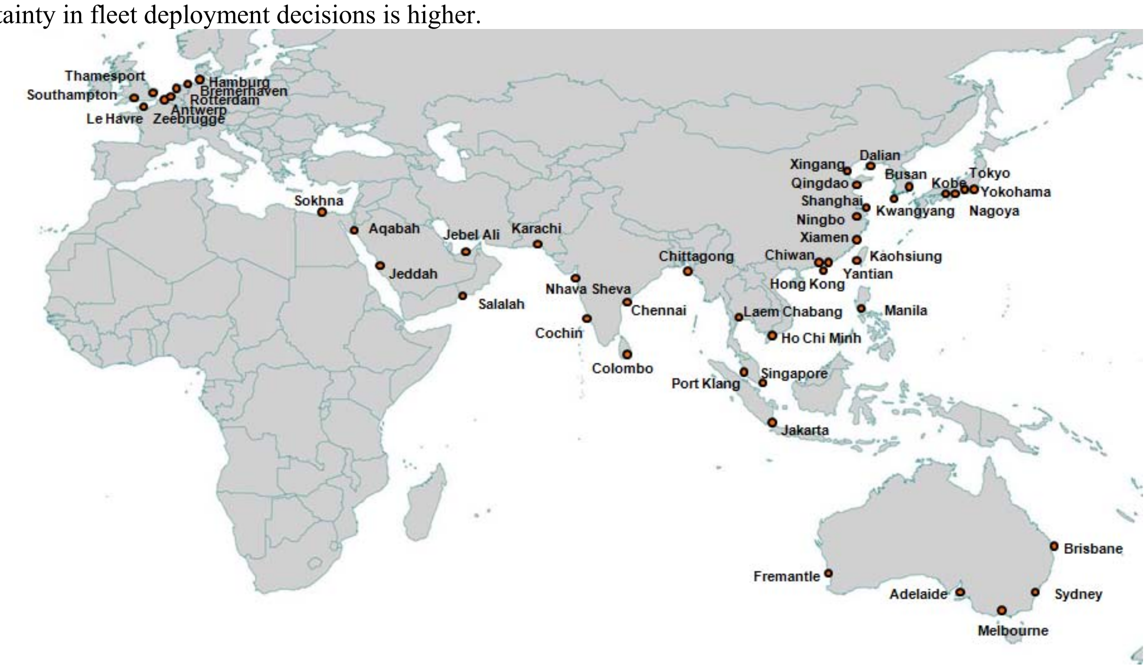


Figure 2. A Liner Shipping Network with 46 Ports

Table 3. Statistical Analysis of Solution Quality

Demand Variability κ	z'_j Average	z'_j Std. Dev.	z_N^k Average	z_N^k Std. Dev.	95% Confidence Interval $\overline{L}$	95% Confidence Interval $\overline{U}$	Relative Optimality Gap
0.2	26.58	0.201	26.59	0.035	26.58	26.60	0.1%
0.5	26.58	0.506	26.60	0.086	26.51	26.64	0.5%
0.9	26.26	0.988	26.28	0.162	26.12	26.35	0.9%

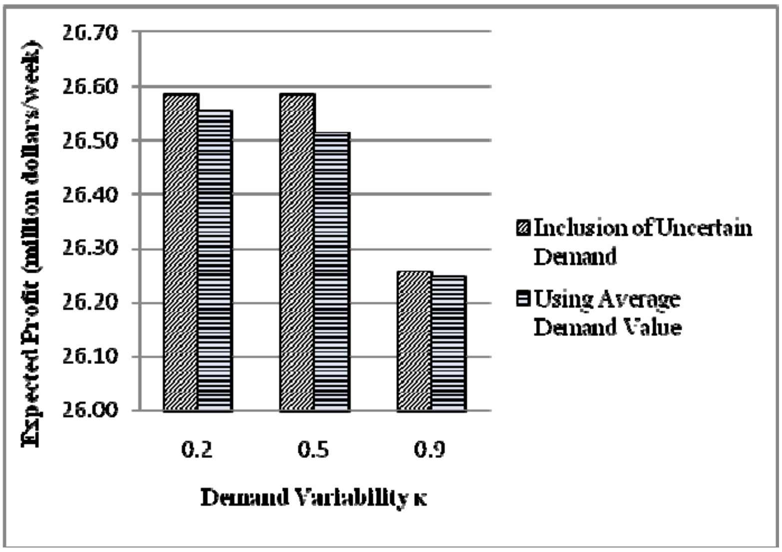


Figure 3. Comparison of Expected Profit with and without Considering Demand Uncertainty

6. CONCLUSIONS

In this paper, we have investigated a novel LSFDP arising in the liner shipping industry. Many practical constraints in liner shipping operations, such as container handling time, weekly service frequency, transshipment, multiple routing plans, are incorporated. In view of the uncertain demand, we optimize the fleet deployment decisions by maximizing the expected profit. The case study based on realistic problems shows that the problem can be solved efficiently with the sample average approximation method and the statistical relative optimality gap is very tight. Results also suggest that the inclusion of uncertain demand into the model is preferable compared with using the deterministic average demand value.

ACKNOWLEDGEMENTS

This study is supported by the research grant WBS No. R-264-000-244-720 from the NOL Fellowship Programme.

7. REFERENCES

Cho, S.C. and Perakis, A.N. (1996). Optimal Liner Fleet Routeing Strategies. *Maritime Policy and Management 23(3)*: 249-259.

Gelareh, S. and Meng, Q. (2010). A Novel Modeling Approach for the Fleet Deployment Problem within a Short-Term Planning Horizon. *Transportation Research 46E*: 76-89.

Jaramillo, D. I., and Perakis, A.N. (1991). Fleet Deployment Optimization for Liner Shipping Part 2. Implementation and Results. *Maritime Policy and Management, 18 (4)*: 235-262.

Lp_solve (2010). Lp_solve5.5.0.15 Reference Guide. http://lpsolve.sourceforge.net/5.5/.

Mak, W.K., Morton, D.P. and Wood, R.K. (1999). Monte Carlo Bounding Techniques for Determining Solution Quality in Stochastic Programs. *Operations Research Letters 24*: 47-56.

Meng, Q. and Wang, T. (2010). A Chance-Constrained Programming Model for Short-Term Liner Ship Fleet Planning Problems. *Maritime Policy and Management*, In Press.

Mourão, M.C., Pato, M.V. and Paixão, A.C. (2001). Ship Assignment with Hub and Spoke Constraints. *Maritime Policy and Management, 29 (2)*: 135-150.

Perakis, A.N. and Jaramillo, D. I. (1991). Fleet Deployment Optimization for Liner Shipping Part 1. Background, Problem Formulation and Solution Approaches. *Maritime Policy and Management, 18 (3)*: 183-200.

Powell, B.J. and Perakis, A.N. (1997). Fleet Deployment Optimization for Liner Shipping: an Integer Programming Model. *Maritime Policy and Management 24(2)*, 183-192.

Shapiro, A. (1991). Asymptotic Analysis of Stochastic Programs. *Annals of Operations Research 30*. 169-186.

Verweij, B., Ahmed, S., Kleywegt, A.J., Nemhauser, G. And Shapiro, A. (2003). The Sample Average Approximation Method Applied to Stochastic Routing Problems: a Computational Study. *Computational Optimization and Applications 24*: 289-333.

A SOLUTION ALGORITHM FOR GREEN FUEL MANAGEMENT OF A SHIP

Hwa-Joong Kim[1]*, Kwang-Tae Kim[2], Dong-Sik Park[2], Hyo-Jeong Kim[2], and Young-Tae Chang[2]

Graduate School of Logistics
Inha University
Incheon 402-751
KOREA
e-mail: {hwa-joong.kim@inha.ac.kr, syndicate@inhaian.net, pdsfgg@inha.edu, namu220@inha.edu, ytchang@inha.ac.kr}

Abstract: Greenhouse gas emissions from shipping are an increasing concern so that many shipping companies are making efforts to reduce the speed of ships aiming at reducing greenhouse gas emissions. Motivated from this practice, this paper aims at finding the optimal ship speed, fueling ports and fueling amount at the ports for a given single ship route. The objective of the problem is to minimize fixed fuel ordering, fuel purchase, fuel inventory holding, and ship time costs as well as environmental cost related to greenhouse gas emissions. We suggest an optimal algorithm by extending the shortest path algorithm. We demonstrate the procedure of the algorithm using sample data.

1. INTRODUCTION

Compared with the other transport modes such as road and aviation, maritime transport has been able to dodge international environmental regulations for a long time due to its image as an environmental-friendly mode. However, over the past half century, maritime transport has been burdened with rapid growth of international trade. On account of this situation, maritime transport is now regarded as a main contributor to harming the environment by emitting the greenhouse gases. Maritime transport is the second largest CO_2-emissions contributor in transportation sector as can be seen in Figure 1.

Following the Kyoto Protocol and UN's sequential actions, the International Maritime Organization (IMO) has been discussing various measures to mitigate greenhouse gas emissions from international shipping. In addition, the Intergovernmental Panel on Climate Change suggests implementing emission trading scheme (ETS) for international shipping as a cost effective manner of compliance with emission regulation. Historically, ETS has significance as the cornerstone of the European Union's strategy as well as the first international emission trading system for dealing with the climate change. To respond to these recent developments, many shipping companies have used technological and operational approaches such as speed reduction, power and propulsion systems, vessel hull design, voyage optimization, renewable energy, and so forth to save fuel consumption generating the greenhouse gas emissions (International Maritime Organization, 2009). Among the methods, ship speed reduction is a simple and effective way for reducing greenhouse gas emissions. A recent news release by the website of Oceana states 'Reducing commercial ship speeds, by only a few knot, yields salutary results for both shipping companies and the environment. "Slower steaming" saves on fuel consumption and cost while also releasing fewer greenhouse gas emissions. The commercial shipping industry is

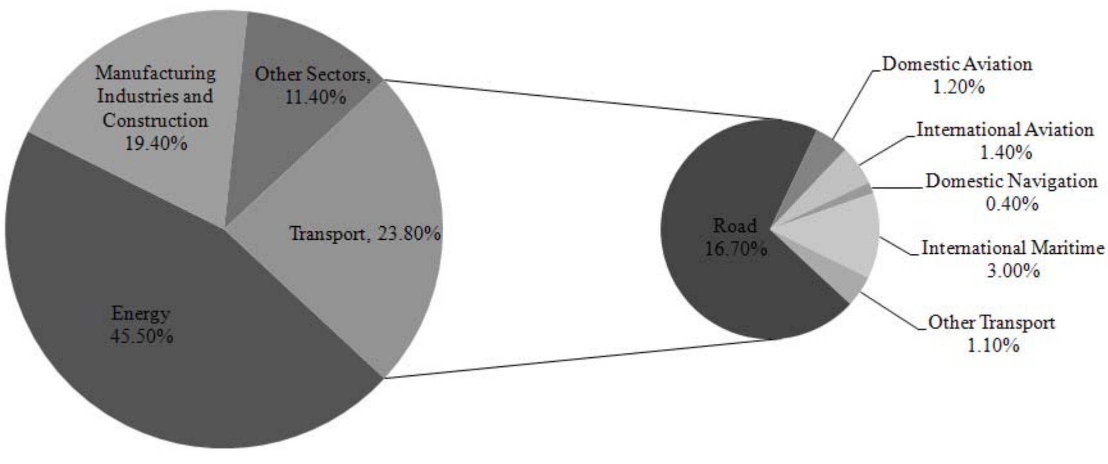


Figure 1. 2006 World CO2 Emissions from Fossil Fuel Combustion (Source: OECD (2009))

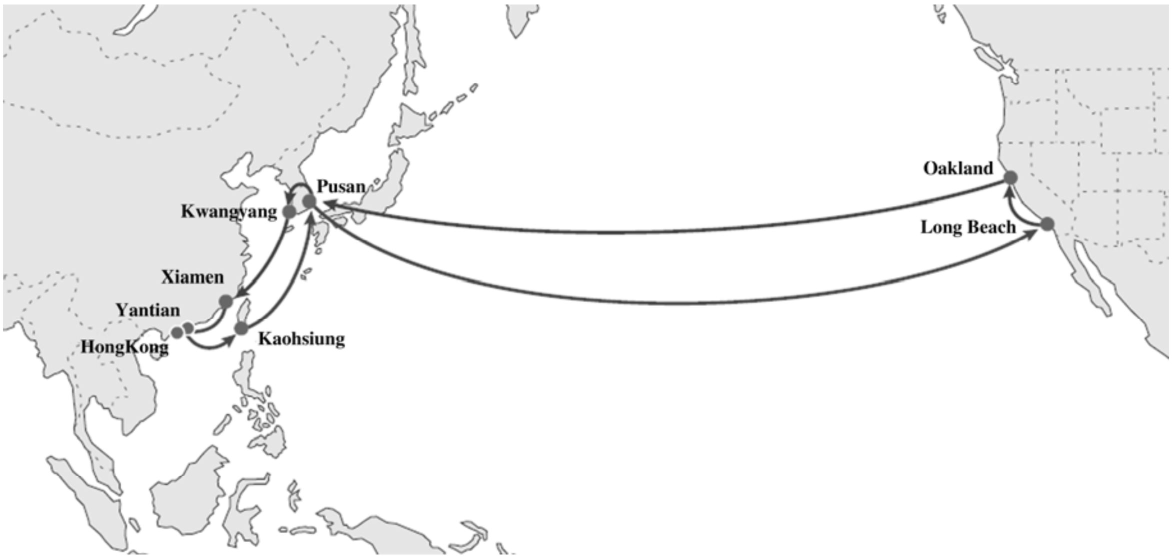


Figure 2. Example of a Shipping Route

beginning to catch on to the benefits of slower steaming and many companies are now enjoying the rewards of more sustainable emissions through monetary savings.' (Fagerholt *et al.* 2010).

Under this context, the objective of this paper is to optimize the ship speed during the navigation of a ship. Our study applies to an exemplary fixed shipping route, such as the one depicted in Figure 2 with eight ports calls. Here, slower steaming results in lower fuel consumption, smaller number of fueling ports, and lower CO_2 emissions, while it results in more navigation time, and vice versa. By taking account of the trade-offs, we deal with the problem of determining fueling ports, fueling amount, and the ship speed for a given shipping route of a single ship. In fact, our problem is a variant of the classical fuel management problem, which determines fueling depots and amount, in that our problem determines the ship speed by considering CO_2 emissions. We suggest an optimal algorithm by extending the shortest path algorithm on a directed acyclic graph to solve the problem.

The remainder of this paper is organized as follows. The next section summarizes previous research related to our problem. In Section 3, we describe the problem details and define related cost functions. Section 4 describes the optimal algorithm and Section 5 presents numerical experiments using a sample data. The final section describes the summary and suggestions for future research.

2. LITERATURE REVIEW

The relevant literature to our problem can be overall twofold: the fuel management problem and the problem of determining the ship speed. For fuel management research, there is no research article in the field of maritime industry to the best of our knowledge, though marine fuel management is one of the most important operational problems. Therefore, we review previous research on the aircraft fuel management problem with some articles. On the other hand, several researchers have used the speed reduction of ships as an option for saving the fuel consumption and thereby for alleviating harm to the environment. For convenience of the literature review, we classify previous research articles on determining ship speed with respect to their research objectives: fuel consumption saving and reduction of greenhouse gas emissions.

For the fuel management problem, Darnell and Loflin (1977) considered the problem of determining the fueling station and vendor for each flight by considering prices, availability, fuel consumption, flight data, and cost of tankerage as the objective function and constraints. To solve the problem, they used a linear programming model implemented at National Airlines. They reported that National Airlines saved the multi-million dollar using the model. Stroup and Wollmer (1992) considered the fuel cost minimization problem with station and supplier constraints to estimate the potential profit in McDonnell Doulas. The station and supplier constraints are an upper and lower bound on the amount of fuel that may be purchased at a particular station and supplier, respectively. The problem is formulated as a linear programming model and they reported that 5 to 6 % cost savings were achieved using the model. Zouein *et al*. (2002) considered a multi-period capacitated problem and developed a linear programming model. They illustrated the applicability of this model to Middle East Airline's operations and argued to obtain 10 % dollar savings by applying it over a one week planning horizon.

For the problem of fuel consumption saving, Ronen (1982) is a pioneer in the research related to ship speed determination. Using the relation that fuel consumption is a cubic function of speed, he suggested cost models for determining an optimal ship speed for different types of legs in order to cope with high fuel price. Later, Ronen (2010) extended his previous models by considering the weekly shipping service prevalent in maritime industry and suggests a procedure to determine the optimal speed and the number of ships deployed in a route. Brown *et al.* (2007) suggested a linear programming model to optimize the fuel consumption when transit time of a ship is given. Notteboom and Vernimmen (2009) consider the impact of increased bunker costs on the design of liner services and suggested a cost model to analyze the impact of bunker cost changes on the operational costs of shipping firms. Fagerholt *et al.* (2010) considered the problem with constraints of time window for the start of service. They formulated the problem as a non-linear programming model which can be solved by a non-linear programming solver. Then, they proposed an alternative solution methodology, in which the arrival times are discretized and the problem is solved as a shortest path problem on a directed acyclic graph. Extensive computational results showed the superiority of the shortest path approach and the potential for fuel savings on shipping routes.

For the problem of reducing greenhouse gas emissions, some research papers have recently emerged in the literature. Matsukura *et al.* (2009) considered a ship route design problem and suggested a mixed integer programming model with the objective of minimizing CO_2 emission while satisfying customer demand and time window constraints. Using the model and data in the container transportation network in Japan, they obtained the routes of ships with different capacities, speeds, and main engine horsepower. Corbett *et al.* (2009) and Wang *et al.* (2010) developed a profit-maximization equation to estimate route-specific, economically-efficient speeds. They explored policy impacts of a fuel tax and speed reduction mandate on CO_2 emissions.

From the above literature review, it is clear that no previous studies have been undertaken yet to consider the maritime fuel management problem of determining the ship speed by considering CO_2 emissions. This study, therefore, intends to fill the gap in the literature.

3. PROBLEM DESCRIPTION

We summarize first the notations used throughout the paper as follows.

Parameters

- c unit time cost of the ship
- c_{ij} total cost occurred during the navigation from port i to port j
- d_{ij} sailing distance from port i to port j
- e CO_2 emission per unit fuel consumed
- f fuel consumption at designed speed of the ship
- h_i unit inventory holding cost of fuel if the fuel is purchased in port i
- p_i unit fuel purchase cost at port i
- $q_{\max}$ capacity of fuel tanker of ship
- s_i fixed ordering cost charged if fueling at port i
- t tax per unit CO_2 emission
- v designed speed of the ship
- $v_{\min}$ minimum ship speed
- $v_{\max}$ maximum ship speed
- Δ discrete speed interval

Decision variables

- F fuel consumption rate on the voyage depending on ship speed
- Q_{ij} fueling amount at port i for sailing from port i to j
- T_{ij} sailing time from port i to port j
- V ship speed

Fuel consumption can be approximated by a well-known cubic function of speed as follows (Ronen 1982, 2010).

$$F = \frac{f}{v^3} \cdot V^3 \tag{1}$$

For simplicity, we assume that the vessel navigates with the same speed along the route and no time is spent in the ports without loss of generality. For practical reasons, we assume that the ship speed has its maximum speed, $v_{\max}$, and minimum speed, $v_{\min}$. We also assume that the ship speed increases or decreases by a discrete interval value Δ. This assumption may be acceptable in practice since shipping companies make their policy on the ship speed with a discrete number, e.g., 20 knot, 21 knot.

Figure 3 illustrates our problem along a given single ship route with four ports calls, indexed *i, j, k,* and *l* in the figure. Horizontal and vertical axes represent sailing time of the ship and fuel inventory level in the ship, respectively. Q_{ij} represents fueling amount at port i for sailing from port i to j and T_{ij} represents sailing time of the ship from port i to j. Fueling amount Q_{ij} is limited by a fuel tank capacity of the ship q_{max}. After fueling Q_{ij} at port i, the ship consumes fuel at rate F until the fuel is run out at port j, i.e., fuel inventory level becomes zero at port j. Using equation (1) and the transit time-speed formula $T_{ij} = d_{ij}/V$, the fueling amount can be expressed as follows.

$$Q_{ij} = T_{ij} \cdot F = \frac{f}{v^3} \cdot d_{ij} \cdot V^2$$

For determining optimal fueling strategy, we need to consider fixed ordering, fuel purchase, and fuel inventory holding costs as in ordinary inventory control problems, as well as ship sailing time cost. In addition, we need to take account of environmental cost in the determination in order to incorporate CO_2 emitted by the ship.

The fixed ordering cost s_i is associated with ordering fuel at port i. The ordering cost associated with typing and negotiating the purchase order does not depend on the amount of fuel ordered. The fuel purchase cost is the cost of buying the fuel at port i to sail from port i to j, which is expressed as its unit cost multiplied by the amount purchased as follows.

$$p_i \cdot Q_{ij} = p_i \cdot \frac{f}{v^3} \cdot d_{ij} \cdot V^2$$

The inventory holding cost is associated with keeping the fuel in the ship in inventory for a period of time. Typically, when the fuel is carried in inventory, the capital invested is not available for other purposes. That is, this represents a cost of forgone opportunities for other purposes, which is assigned to inventory as an opportunity cost. The fuel inventory holding cost during the navigation from port i to j is calculated by the unit inventory holding cost multiplied by the average inventory from port i to j as follows.

$$h_i \cdot T_{ij} \cdot \frac{Q_{ij}}{2} = h_i \cdot \frac{f}{v^3} \cdot d_{ij}^{\,2} \cdot \frac{V}{2}$$

The ship sailing time cost includes the fixed cost including the capital cost, the maintenance, crew wage, and so forth. The cost from port i to j is represented by the unit cost of ship multiplied by the sailing time as follows.

$$c \cdot T_{ij} = c \cdot \frac{d_{ij}}{V}$$

The environmental cost related to CO_2 emission of the ship is associated with possible tax imposed on CO_2 emission generated from fuel consumption. The cost is calculated by considering tax on unit CO_2 emission, CO_2 emission generated from unit fuel amount, and fuel amount purchased, denoted as

$$t \cdot e \cdot Q_{ij} = t \cdot e \cdot \frac{f}{v^3} \cdot d_{ij} \cdot V^2 .$$

Based on above cost functions, the total cost during the navigation from port i to j is then

$$c_{ij} = s_i + p_i \cdot \frac{f}{v^3} \cdot d_{ij} \cdot V^2 + h_i \cdot \frac{f}{v^3} \cdot d_{ij}^{\,2} \cdot \frac{V}{2} + c \cdot \frac{d_{ij}}{V} + t \cdot e \cdot \frac{f}{v^3} \cdot d_{ij} \cdot V^2 . \qquad (2)$$

Now, the considered problem can be defined as the problem of determining the ship speed, fueling ports, and fueling amount at the ports along the route for the objective of minimizing the sum of fixed ordering, fuel purchase, fuel inventory holding, ship sailing time, and environmental costs. The problem is constrained by minimum and maximum speeds, fuel tank capacity, and speed change by a discrete interval. Finally, the other assumptions made in this research are summarized as follows: (a) no time is needed for fueling at all ports; (b) container weights on the ship does not influence on fuel consumption rate; and (c) the ship is perfect in state, i.e., it is not out of order during its navigation.

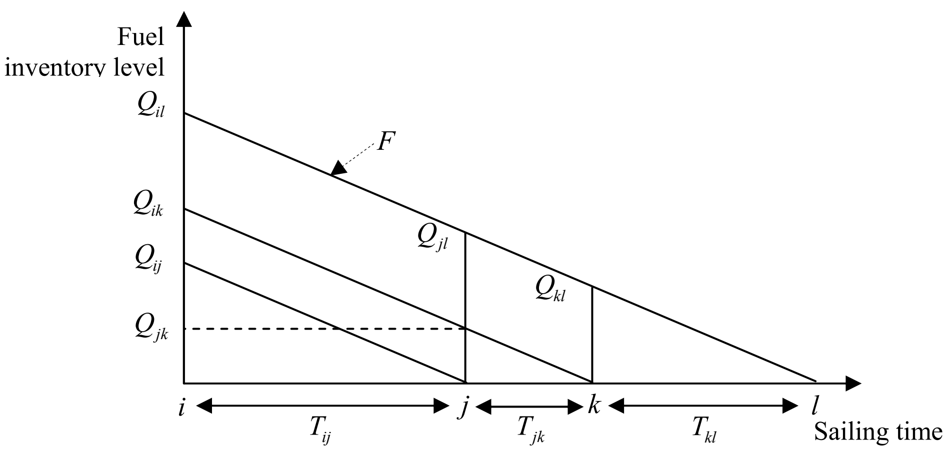


Figure 3. Fuel Inventory Level for Sailing

4. SOULUTION ALGORITHM

The problem considered in this research can be formulated as a shortest path problem under the condition that ship speed V is given. Since ship speed is given, we only need to determine fueling ports. Here, fueling amount is automatically calculated since the amount at a port is exactly the amount that the ship can navigate until the next fueling port. To formulate the problem as a shortest path problem, consider a directed acyclic graph with nodes {1, 2, …, n} representing ports and arcs (i, j) for all $i < j$. The cost of arc (i, j) is the cost of fueling at port i and navigating until port j, i.e., c_{ij} defined in (2). Clearly, the optimal solution of our problem with a given ship speed V is the least cost path from node 1 to node n. Here, the fuel tank capacity restriction can be considered by eliminating the corresponding arcs from consideration, i.e., if the fueling amount Q_{ij} for arc (i, j) is more than the fuel tank capacity $q_{\max}$, the corresponding arcs are eliminated from further consideration. Figure 4 shows an example of the shortest path problem with four ports calls.

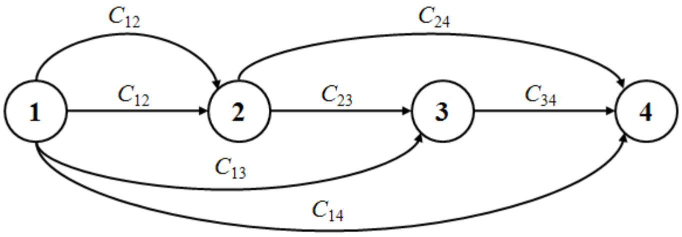


Figure 4. Shortest Path for the Problem with a Given Speed

The shortest path problem can be formulated as a dynamic programming model. Let $H(k)$ be the cost of the cheapest path from node 1 to k, i.e., the optimal cost of the subproblem from node 1 to k, called kth subproblem hereafter, which is defined as

$$H(k) = \min_{1 \le i \le k-1} \{H(i-1) + c_{ik} \mid \frac{f}{v^3} \cdot d_{ik} \cdot V^2 \le q_{\max}\}$$

with $H(1) = 0$. The optimal cost of the problem with a given ship speed $H(n)$ can be obtained by computing $H(k)$ in forward direction, starting from $k = 2$ and ending at $k = n$. Working back gives a corresponding optimal solution.

Then, the original problem can be solved optimally by increasing the ship speed by the discrete speed interval Δ from $v_{\min}$ to $v_{\max}$. The procedure of the optimal algorithm is summarized below.

Procedure

Step 1. Set $V = v_{\min}$

Step 2. Set $k = 2$ and $H(1) = 0$.

Step 3. From $i = 1$ to $k – 1$ satisfying $f/v^3 \cdot d_{ik} \cdot V^2 \le q_{\max}$, calculate

$$H_{ik} = H(i-1) + s_i + p_i \cdot \frac{f}{v^3} \cdot d_{ik} \cdot V^2 + h_i \cdot \frac{f}{v^3} \cdot d_{ik}{}^2 \cdot \frac{V}{2} + c \cdot \frac{d_{ik}}{V} + t \cdot e \cdot \frac{f}{v^3} \cdot d_{ik} \cdot V^2$$

Step 4. Obtain the cheapest path from node 1 to k using

$$H(k) = \min_{1 \le i \le k-1} H_{ik}$$

Step 5. Set $k = k + 1$. If $k \le n$ (index for the last node), go to Step 2 and go to Step 6, otherwise.

Step 6. Set $V = V + \Delta$. If $V \le v_{\max}$, go to Step 2 and stop, otherwise.

The optimal algorithm is a pseudo polynomial algorithm with $O((v_{\max}\text{-}v_{\min})/\Delta \cdot n^2)$ since Steps 2 to 5 can be computed in $O(n^2)$ and the same steps iterate $O((v_{\max}\text{-}v_{\min})/\Delta)$ times in the above procedure.

5. NUMERICAL EXAMPLE

Table 1. Distance Between Ports

unit: nautical miles

	Kwang yang	Xiamen	Yantian	Hong Kong	Kaoh siung	Pusan	Long Beach	Oakland	Pusan
Pusan	103	1023	1292	1333	1676	2607	8531	8895	13873
Kwangyang	-	920	1189	1230	1573	2504	8428	8792	13770
Xiamen	-	-	269	310	653	1584	7508	7872	12850
Yantian	-	-	-	41	384	1315	7239	7603	12581
Hong Kong	-	-	-	-	343	1274	7198	7562	12540
Kaohsiung	-	-	-	-	-	931	6855	7219	12197
Pusan	-	-	-	-	-	-	5924	6288	11266
Long Beach	-	-	-	-	-	-	-	364	5342
Oakland	-	-	-	-	-	-	-	-	4978

In this section, we demonstrate the above procedure using the data collected from various sources. We used the shipping route in Figure 2. In the route, Pusan port is regarded as the starting and last ports, i.e., the ship's sailing sequence is Pusan → Kwangyang → Xiamen → Yantian → HongKong → Kaohsiung → Pusan → Long Beach → Oakland → Pusan, which are in the shortest path network numbered as 1, 2, …, 10, respectively. The ship used in the test is a 5000 TEU container ship with fuel tank capacity 10,681 m^3 (Kim 2006), designed speed 24 knot, and fuel consumption rate at the speed 149 ton per day (Notteboom and Vernimmen 2009). The minimum and maximum speeds of the ship were set to 6 knot and 26 knot respectively and the discrete speed interval was set to 0.1 knot. The distance data between ports is obtained from SeaRates' website, summarized in Table 1. Fuel price (USD per ton) at Pusan, Kwangyang, Xiamen, Yantian, Hong Kong, Kaohsiung, Long Beach, and Oakland, obtained from the websites of Bunker Index, Bunker World, and Shipping Online, were set to 450.0, 450.0, 470.5, 461.5, 438.0, 484.0, 435.5 and 442.0, respectively. Among them, Yantian and Oakland ports' fuel prices were replaced by that of GuangZhou and San Francisco, respectively, due to their absence. The fixed ordering cost at each port was obtained from 10 times fuel price at the port and inventory holding cost at each port was obtained by 0.01 times fuel price at the port. CO_2 emission was set to 3 tCO_2 per fuel ton obtained from the websites of Korea Energy Management Corporation and the emission tax was set to 100 USD per tCO_2 originated from Corbett *et al.* (2009). Finally, the unit time cost of the ship was set to 25,116 USD per day, approximated from the data in Kim (2005).

Figure 5 shows the total costs at different ship speeds. The optimal ship speed and total cost are 9.6 knot and 2,598,965.8 USD, respectively. As expected, the total cost decreases at first and increases after the optimal speed, as the ship speed increases. This can be explained by the trade-off relationship between the ship time cost and the other costs, as described earlier. In fact, the optimal ship speed 9.6 knot is significantly lower than the speed used in practice since our model did not include the daily value of containers due to its absence. Once the value is incorporated into our model, the optimal ship speed will be increased. In order to exactly incorporate the daily value of containers into our model, we

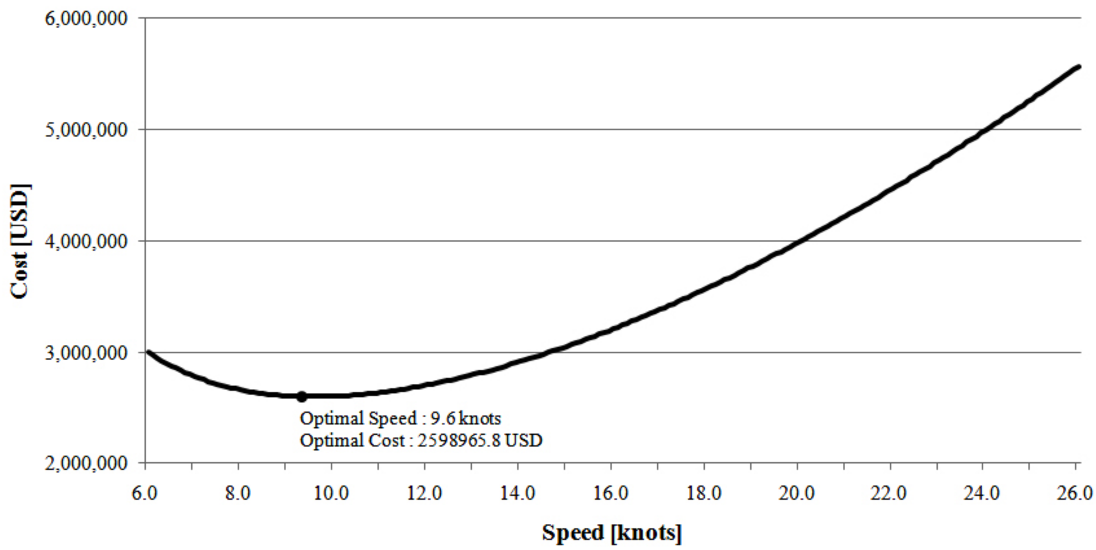


Figure 5. Total Cost Function

Table 2. Optimal Solution of Each Subproblem at the Optimal Speed

Last fueling port (i)	kth subproblem								
	2	3	4	5	6	7	8	9	10
1	**19233.2***	**159957.7***	204205.9	211073.3	269802.1	440710.8	1922004.4	2035217.7	3841374.6
2	-	162619.9	206330.7	213116.2	271159.9	440208.7	1909668.5	2022154.6	3818367.4
3	-	-	**203833.9***	**209933.1***	**262293.5***	416440.4	1809017.6	1917790.8	3674973.7
4	-	-	-	214308.9	264642.3	413057.4	1761286.8	1866891.2	3575579.9
5	-	-	-	-	263973.8	409962.1	1722184.9	1824417.7	3473531.7
6	-	-	-	-	-	**409724.8***	1740586.7	1846233.7	3568398.9
7	-	-	-	-	-	-	**1595994.4***	1690803.4	3245267.1
8	-	-	-	-	-	-	-	**1653077.9***	2623746.1
9	-	-	-	-	-	-	-	-	**2598965.8***
Optimal last fueling port	**1** (Pusan)	**1** (Pusan)	**3** (Xiamen)	**3** (Xiamen)	**3** (Xiamen)	**6** (Kaohsiung)	**7** (Pusan)	**8** (Long Beach)	**9** (Oakland)
Optimal fueling amount (ton)	4.3	42.3	53.5	55.2	69.3	107.9	353.0	368.0	574.0

* minimum cost of the kth subproblem

need to consider the values varying during the ship's navigation since containers are loaded and unloaded at ports, which may result in the modification of our model. On the other hand, we describe how to derive the optimal cost at the optimal speed level in Table 2, which shows the optimal solution of each subproblem. The numbers in the upper right part of the table are the solution values of candidates in subproblems, e.g., 416440.4 is the solution value when fueling in port 3 in the 7th subproblem. As can be seen in the table, the optimal fueling ports are ports of Pusan, Xiamen, Kaohsiung, Pusan, Long Beach, and Oakland, and the optimal fueling amouts (ton) at the ports are 42.3, 69.3, 107.9, 353.0, 368.0, and 574.0 in the above data setting.

6. CONCLUSION

We considered the fuel management problem of a ship, which determines the ship speed, fueling ports, and fueling amount at the ports for the objective of minimizing the sum of fixed ordering, fuel purchase, fuel inventory holding, ship sailing time, and environmental costs. The environmental cost was considered as possible tax imposed on CO_2 emission of the ship during its navigation. To solve the problem, we suggested a pseudo polynomial algorithm by extending the shortest path algorithm after defining the cost factors. A numerical experiment on sample data showed that the optimal ship speed is 9.6 knot.

This research can be extended in various ways. At first, although the sensitivity analysis was not performed in this paper due to space limitation, it is worthwhile analyzing the sensitivity of tax, fuel price, ship daily cost, and so forth. Second, as described earlier, the container value varying over the route needs to be considered in future research and the ship speed may be increased. Third, since this research considered a problem with a discrete speed interval, it is meaningful to consider a continuous ship speed problem. Third, a ship schedule may not be easily changed in real practice and hence, we need to consider a time window constraint of the ship service time at the ports. Fourth, the ship speed may be different in different legs of the route in practice and therefore, the problem with different ship speeds in legs is needed to consider as a future research topic.

7. REFERENCES

Brown, G. G., Kline, J. E., Rosenthal, R. E., and Washburn, A. R. (2007). Steaming on Convex Hulls, *Interfaces, 34:* 342-352.

Bunker Index: Bunker Index Home Page. http://www.bunkerindex.com.

Bunker World: Bunker World Home Page. http://www.bunkerworld.com.

Corbett, J. J., Wang, H., and Winebrake, J. J. (2009). The Effectiveness and Costs of Speed Reductions on Emissions from International Shipping, *Transportation Research Part D, 14:* 593-598.

Darnell, D. W., and Loflin, C. (1977). National Airlines Fuel Management and Allocation Model, *Interfaces, 7:* 1-16.

Fagerholt, K., Laporte, G., and Norstad, I. (2010). Reducing Fuel Emissions by Optimizing Speed on Shipping Routes. *Journal of the Operational Research Society, 61:* 523-529.

Hyundai Merchant Marine: Hyundai Merchant Marine Home Page. http://www.hmm21.com.

International Maritime Organization (2009). Second IMO GHG Study 2009.

Korea Energy Management Corporation: Korea Energy Management Corporation Home Page. http://www.kemco.or.kr.

Kim, H. C. (2006). *Korean Ship*. Jisungsa, Seoul, Korea.

Kim, T. W. (2005). Optimal Containership Size by Total Shipping Cost Analysis. Master Thesis, Korea Maritime University, Korea.

Matsukura, H., Udommanhuntisuk, M., Yamato, H., and Dinariyana, A.A.B. (2010). Estimation of CO_2 Reduction for Japanese Domestic Container Transportation based on Mathematical Models. *Journal of Marine Science and Technology, 15:* 34-43.

Notteboom, T. E., and Vernimmen, B. (2009). The Effect of High Fuel Costs on Liner Service Configuration in Container Shipping. *Journal of Transport Geography, 17:* 325-337.

Oceana: Oceana Home Page. http:// www.oceana.org.

OECD (2009). Reducing transport GHG emissions - Opportunities and Costs, *International Transport Forum.*

Ronen, D. (1982). The Effect of Oil Price on the Optimal Speed of Ships. *Journal of the Operational Research Society Research, 33:* 1035-1040.

Ronen, D. (2010). The Effect of Oil Price on Containership Speed and Fleet Size. *Journal of the Operational Research Society*, doi: 10.1057/jors.2009.169.

SeaRates: SeaRates Home Page. http://www.searates.com.

Shipping Online: Shipping Online Home Page. http://www.shippingonline.cn.

Stroup, J. S., and Wollmer, R. D. (1992). A Fuel Management Model for the Airline Industry. *Operations Research, 40:* 229-237.

Wang, H. (2010). Speed Reduction as a Policy Option in Greenhouse Gases Mitigation in the Shipping Industry, *Marine Policy*, doi:10.1016/j.marpol.2009.12.006.

Zoueuin, P. P., Abillama, W. R., and Tohme, E. (2002). A Multiple Period Capacitated Inventory Model for Airline Fuel Management: A Case Study. *Journal of the Operational Research Society, 53:* 379-386.

ALLOCATION OF HANDLING DEVICES FOR BARGES UNLOADING

Nenad Bjelić[1], Vidović Milorad[2], Katarina Vukadinović[3]

[1,2,3]Faculty of Transport and Traffic Engineering,
University of Belgrade,
Vojvode Stepe 305,
Belgrade, Serbia
e-mails: n.bjelic@sf.bg.ac.rs[1], mvidovic@sf.bg.ac.rs[2], k.vukadinovic@ sf.bg.ac.rs[2]

Abstract: Unloading gravel from barges is activity which, if it is not efficiently realized, may have huge impact on overall productivity of the gravel distribution system. This impact is expressed though increased time barges spend waiting unloading, instead transporting gravel from production sites. Therefore, providing efficient and cost effective service of loaded river barges needs appropriate allocation plan for handling equipment. In order to utilize handling facilities efficiently, and to minimize waiting time of barges, it is necessary to consider both: assignment of handling equipment to unloading locations, and the order of servicing different unloading locations.
This paper presents the model that can be used for solving problem of allocating handling devices to the barges in case when tasks appearance times are distributed over planning horizon. among the mixed integer programming formulation which is not appropriate to solve problems of larger dimensions, a ng reedy heuristic algorithm “Double Deep Increment” has been also proposed.
Keywords: gravel distribution, allocation of unloading devices, Mix Integer Programming, heuristics

1. INTRODUCTION

Gravel distribution by inland water transportation includes three main phases: loading of gravel by a suction dredger into barges, transport of gravel to the ports or unloading locations, and unloading of gravel by a handling facility that usually consists of pontoon mounted crane and belt conveyor. Because of high costs, a number of handling facilities is usually relatively small, and requires successive relocation of handling equipment between different unloading locations. Accordingly, providing efficient and cost effective service of loaded river barges needs appropriate allocation plan for handling equipment, which means defining sequence of unloading locations that should be served by each handling device.

The problem may be introduced in following way. For a given collection of barges unloading tasks find a set of assignments to minimize the sum of the service times including waiting for service and handling devices transfer times. The problem of this type may be considered as dynamic handling devices scheduling problem, where tasks service ready times are known after the beginning of the planning interval, or as static problem, where all tasks are already known when the scheduling plan is determined. In this paper, only static problem (SHDAP) is studied.

The objective of this problem is minimization of barges service times so that they can spend as much as possible time in houling goods, i.e. by making profit to an owner. It is important to note that this problem is not only applicable to gravel distribution. It can be applied to other transportation modes, where resources are cars, trucks, containers,...,etc.

It should be noted that although barge transportation in general is present in practice for a long time there are not much paper related to its operational control, ***Christiansen et al. (2007)***. The first published paper related to management of fleet of barges is of ***O’Brien and Crane (1959)*** in which authors considered effects of implementing different scheduling strategies on existing barge and towing boats fleet utilization. A tool used was simulation. In ***Schwartz (1968)*** linear mixed integer problem formulation was used for solving scheduling problem of fleet of barges and towing boats and objective of minimizing related costs. Another paper that uses simulation for modeling barges operational control is ***Taylor et al. (2005)*** in which authors are developed simulation based DSS as a help in dispatchers decision making. Developed DSS is nowadays in use in American Commercial Barge Line, LLC on tasks executed on Ohio River.

Among mentioned papers in literature one can find papers related to the operational control of barges which are used in gravel transportation. Two such papers are ***Vukadinovic and Teodorovic (1994)*** and ***Vukadinovic et al. (1997)*** where authors used fuzzy logic and artificial neural networks in managing systems characterized with stochastic processes.

The problem of allocating of handling devices for barges unloading is studied by ***Vidovic and Vukadinovic (2006)*** who formulated Handling Devices Allocation Problem – HDAP. Authors considered static case of HDAP (SHDAP) and presented two approaches to its formulation. First one is based on three-dimensional assignment problem, while the second is based on similarities of SHDAP with the Static Berth Allocation Problem – SBAP, respecting differences between those two problems. In addition, authors presented three step heuristic algorithm to solving SHDAP (CLASORD – CLustering ASsignment ORDering). First step classifies barges unloading tasks according to its locations by performing K-means clustering technique. Second step determines sequences of task realizations within each cluster and assigns devices to clusters. This is realized by solving two dimensional assignment problem. Third step tries to find solution improvements by changing tasks between clusters.

However, ***Vidovic and Vukadinovic (2006)*** consider only static case of the problem in which all tasks are present in the system at the beginning of planning horizon. Therefore, logical step further could be to consider more general problem, closer to real sysems where not all of tasks are present, i.e. not all of barges arrived at unloading positions at the beginning of planning horizon. Research presented in this paper tries to make this step forward, and to formulate the problem in the case when not all of handling tasks are available at the beginning of observed interval, and to propose solution approach.

Remaining of the paper is organized as follows. Section 2 puts forward the mathematical formulation of the problem. Section 3 presents briefly description of so far developed, well known, heuristics algorithms for the problem solving and gives detailed description of new heuristic approach. In section 4, different solving approachs are compared on test example. Analysis of results, as well as direction for future research, are presented in section 5.

2. PROBLEM FORMULATION

Allthough the berth allocation problem is considered as similar to the HDAP, there are also another related approaches. One important area related to the problem analysed here is operational control of internal transportation system – OCITS (***Van der Meer 2000)***. The objective of this study is construction of tasks sequences that each of internal transportation devices will follow while servicing tasks within planning horizon. The objective is to minimize tasks' service times, where an task realization implies movement from a delivery location of previous task to a pick up location of next, loading item onto a device, movement to delivery location of next task and finally, unloading item from device. Because of similarities with OCITS, it is possible to cnsider SHDAP as a Traveling Repairman Problem with Time Windows - TRPTW.

TRPTW formulation implies that the original network comprising locations of tasks occurrence should be transformed into a network of so called task-nodes. In such network each task is represented as a node, and then the problem is to find route which visits each task-node exactly once by some of m available devices. Distance between two task-nodes, *i* and *j* that here represents completion time of task *j* after the task *i* is completed, can, be calculated by summarizing time needed for moving device *v* from the location of task *i* to the location of task *j* and time needed to complete servicing task *j*, as shown in Figure 1.

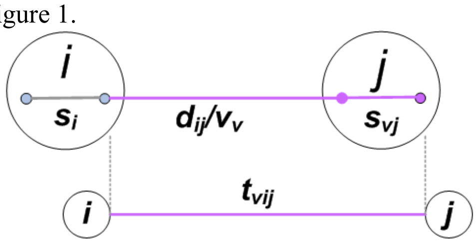


Figure 1. Transformation into a network of task-nodes

If d_{ij} represents distance between locations of tasks *i* and *j*, v_v denotes travel speed of the device *v* and s_{vj} represents service time of task *j* by device *v*, than time distance in transformed network is calculated as $t_{vij} = \frac{d_{ij}}{vv} + s_{vj}$. Because handling devices generally have different characteristics regarding relocating speed and/or unloading capacity as a each device will have its own network. After this transformation, the problem to be solved can be considred as a multiple Traveling Repairman Problem with Time Windows – mTRPTW.

Linear programming formulation of mTRPTW which as a result defines optimal task servicing sequences for all of available devices during planning horizon is formulated as MIP in the following way:

$$\min \sum_{i \in P} (D_i - r_i) \tag{1}$$

s.t.

$$\sum_{v \in V} \sum_{j \in N} x_{ij}^v = 1 \quad \forall i \in P \tag{2}$$

$$\sum_{j \in N} x_{ij}^v - \sum_{j \in N} x_{ji}^v = 0 \quad \forall i \in P, \forall v \in V \tag{3}$$

$$\sum_{j\in P} x_{0j}^{v} \leq 1 \quad \forall v \in V \tag{4}$$

$$\sum_{j\in P} x_{i\,n+1}^{v} \leq 1 \quad \forall v \in V \tag{5}$$

$$D_i + t_{ij}^{v} - D_j \leq M\left(1 - x_{ij}^{v}\right) \quad \forall i, j \in P, \forall v \in V \tag{6}$$

$$D_0^{v} + t_{0j}^{v} - D_j \leq M\left(1 - x_{0j}^{v}\right) \quad \forall j \in P, \forall v \in V \tag{7}$$

$$D_i + t_{i\,n+1}^{v} - D_{n+1}^{v} \leq M\left(1 - x_{i\,n+1}^{v}\right) \quad \forall i \in P, \forall v \in V \tag{8}$$

$$D_i + r_i - s_i^{v} \leq M\left(1 - \sum_{j\in P\cup\{0\}} x_{ji}^{v}\right) \quad \forall i \in P, \forall v \in V \tag{9}$$

$$D_0^{v} = A^{v} \quad \forall v \in V \tag{10}$$

$$D_{n+1}^{v} \geq 0 \quad \forall v \in V \tag{11}$$

$$\sum_{v\in V}\sum_{j\in N} x_{n+1\,j}^{v} = 0 \tag{12}$$

$$\sum_{v\in V}\sum_{j\in N} x_{i0}^{v} = 0 \tag{13}$$

$$x_{ij}^{v} \in \{0,1\} \quad \forall i, j \in N, \forall v \in V \tag{14}$$

Where:
N-set of task-nodes $\{0,...,n+1\}$ to be visited by devices, indexed with *i* and *j*
P- set of task-nodes $\{1,...,n\}$ without nodes representing starting/ending point of every device (depot)
V-set of devices $\{1,...,|V|\}$ for realization of unloading tasks, indexed with *v*, $|V|$ represents cardinality of set *V*
n- number of tasks for execution, $n=|P|$
r_i – appearance time of task-node *i* (left hand side boundary of time window)
t^{v}_{ij} – time needed by device *v* to come from task-node *i* to task-node *j*. $\forall v \in V$, $\forall i, j \in N \wedge i \neq j$
s^{v}_{i} – service time of task *i* by device *v*, $\forall v \in V$, $\forall i \in P$
A^{v} – time when device *v* becomes available for task execution, $\forall v \in V$
M – big number
D_i – completition time of task *i*, $\forall i \in P$
D^{v}_{0} – depot (task-node 0) leaving time of device *v*, $\forall v \in V$
D^{v}_{n+1} – returning time of device *v* into the depot (task-node *n+1*), $\forall v \in V$

$$x_{ij}^{v} = \begin{cases} 1 - if\ device\ v\ is\ moving\ from\ task\ i\ to\ task\ j \\ 0 - otherwise \end{cases}, \ \forall v \in V\ ,\ \forall i, j \in N$$

Objective function (1) minimizes sum of tasks' waiting times, until the end of service, i.e. it minimizes sum of total times between a task's appearance and competition. Sets of constraints (2)-(5) and (12)-(13) provide realization of flows trough the network. Constraints (2) ensures that each task-node will be visited exactly once; constraints (3) is flow conservation.; Constraints (4) give possibility to devices to leave task 0 (depot), and constraints (13) do not allow them to return to it; similarly, constraints (5) and (14) provide that devices can not leave task n+1, but that they can return to it.

Constraints (6) – (8) define compatibility requirements between routes and schedules, while expressions (9)-(11) repesent time window constraints. Constraints (9) define left hand side boundary of the window, i.e. time after which task service is possible, while right hand side boundary does not exist.

Nature of used decision variables, x^{v}_{ij} (binary values) and D_i (real values), as well as their mutual relationship, put this problem into a class of Mixed Integer Problems - MIP. Since mixed integer problems require significant amount of compytatioal, problem size is reduced to relatively small number of handling tasks. For example, CPLEX 12, run on AMD Phenom II 2.61 GHz processor with 2 GB of RAM, encountered memory problems related with the size of branch-and-cut tree even for the problem of seventeen tasks. Therefore, necessity for efficient heuristic algorithm for solving SHDAP becomess obvious.

3. HEURISTIC ALGORITHMS

Transformation of original network into task-node form and then giving opportunity of solving the problem as a mTRPTW has additional advantage in posibility of implementing some of well known heuristic approaches developed for solving TSP. One good example is possibility of insertion heuristics aoolication. This greedy heuristics, in each iteration, builds solution by inserting a considered task into a sequence of already allocated. Device to which task is going to be allocated, as well as location where task is inserted, are selected under criterion of minimal increasing of the objective function. More detailed escription of the heuristics can be found in ***van der Meer (2000)***.

Although in above mentioned study, insertion heuristics provided satisfactory results, when it is applied to the HDAP, results cannot be evaluated as satisfying as is concluded in ***Bjelic (2009).*** Therefore, this paper proposes new *Double Deep Increment* - *DDI* heuristics algorithm to solving HDAP.

Double Deep Increment heuristics also belongs to a class of greedy heuristic algorithms with the number of iterations which is equal to number of tasks to be allocated. Heuristics is based on the computation of averaged performance of asigning tasks to the devices, where this performace is determined in each iteration. Average performance AV_{ik} is calculated for every combination of unallocated tasks $k \in B$ and available devices $i \in H$. Assignment is done for the task – device combination with the lowest value of performance. Assigned task is then added as the lst element of the set of tasks to be served by selected device.

Performances values consider two caracteristic measures. One determined effects of the task assignment, while the another compute effects of the assignment to all other still unallocated tasks. Namely, proposed performance measure estimates effects of assigning task *k* to device *i* by computing:

- time that task *k* is going to spend in system when it is allocated to device *i* as the last task in *i*'s sequence, and
- times that all other unassigned task would spend in the system if task k is going to be allocated to device *i* as the next task in *i*'s sequence.

In this way decisions on te allocation are made in respect to consequences of the assignments, not only by the assignment effects itself. For example, if only first level of assignments was considered, tasks that appear later would have advantage because the time they spend in the system would be equal to serving time, i.e. they would be served without waiting. Cause for this lies in fact that devices would have enough time to come to locations of these tasks and wait for their appearance. On the other hand, because devices usually does not have enough time for moving to locations where tasks appeared at the beginning of planning horizon, those tasks obviously accumulate time spend in the waiting to the beginning of service. Therefore, if decision is made only by considering first performance related to assignment effects, values of performance measurs for tasks appearing later in planning horizon will be better.

However, because waiting for appearance of tasks arriving later, at the same time increases waiting times of earlier appeared tasks And this is why it is so important to include in analysis unassigned tasks as well.

Pseudo code of propsed Double Deep Increment heuristics is:

Set $r_i = \{0\}$, *all devices* $\forall i \in H$ *are located at node* 0

Set $j=0$ to be the last element of set r_i

Set $ET_{i0} = 0$

while $B \neq \varnothing$

for all $i \in H$

for all $k \in B$

/* *calculation of completion time* ET_{ik} *of task k, when is served by device i after task j* */

$$ET_{ik} = ET_{ij} + \max(A_k; \frac{d_{jk}}{v_i}) + \frac{q_k}{p_i}$$

for all $l \in B \setminus \{k\}$

/* *calculation of completion time* ET_{il} *of task l, when is served by device i after task k* */

$$ET_{il} = ET_{ik} + \max(A_l; \frac{d_{kl}}{v_i}) + \frac{q_l}{p_i}$$

/* *calculation of average performance* AV_{ik} *of asigning task k to the device i* */

$$AV_{ik} = \frac{ET_{ik} - A_k + \sum_{\forall l \in B \setminus \{k\}} (ET_{il} - A_l)}{|B|}$$

/* *assigning the task k to device i* */

Find assignment (i,k) so that $\min_{i,k}(AV_{ik})$

/* *updating sets B, and* r_i */

$B = B \setminus \{k\}$

$r_i = r_i \cup \{k\}$

$j = k$

end while

Where:

H – set of handling devices

B – set of handling tasks, which elements are in the same time nodes where tasks are located, hence $B \cup \{0\}$ represents all networks node including depot located in node 0

$|B|$ - cardinality of set B

$r_i \subseteq B, \forall i \in H$ – set of tasks served by handling device *i* ordered by the service order (first element is the index of the task which is served first)

$\mathbf{A_k}$ – appearence time of task *k*

$\mathbf{d_{jk}}$ – distance between nodes where located tasks $j \in B$ and $k \in B$

$\mathbf{q_k}$ – quantity of load to be handled while serving task *k*

$\mathbf{v_i}$ – travel velocity of the device *i*

$\mathbf{p_i}$ – productivity of the device *i*

4. COMPUTATIONAL EXPERIMENT

In this chapter we co mpared solutions obtained by optimization model, insertion heuristics and proposed Double Deep Increment heuristics. Problems are generated on the following network (Figure 2).

- distances (km) between network nodes which are also tasks locations are shown below

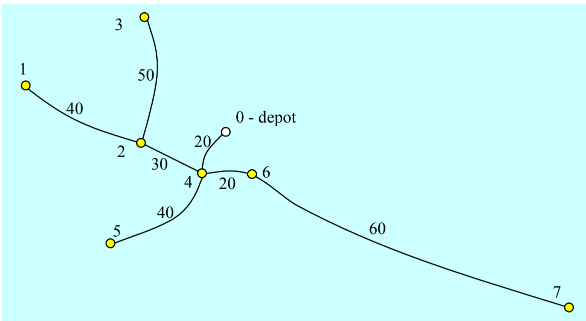


Figure 2. Layout of test network

- there are three unloading devices with productivities of p_1=100t/h, p_2=150t/h, p_3=200t/h
- travel velocities of handling devices are $v_1=v_2=v_3$=5km/h
- task can appears in each location with the same probability
- planning horizon is two days
- time between task arrival is uniform ly distributed within the interval of [5,9] hours for the case of se ven tasks in planning horizon, and within the interval of [3,5.4] for the case of twelve tasks in planning horizon
- each task represents barge loaded by 1000t of gravel
- at the beginning of planning horizon all hndling devices start from depot
- five problem instances, shown in Table 1. and Table 2., with diggeent task appearance times and locations are analyzed both for cases of seven and twelve tasks in planning horizon

Table 1. Parameters of five problem instances in case of seven tasks in planning horizon

Task No.	Instance 1		Instance 2		Instance 3		Instance4		Instance 5	
	Appearance time [h]	Location	Appearance time [h]	Location	Appearance time [h]	Location	Appearance time [h]	Location	Appearance time [h]	Location
1	6.474	6	7.878	4	6.962	3	8.241	2	8.613	6
2	12.691	5	14.778	6	14.091	2	15.054	5	15.972	2
3	17.987	2	21.942	3	21.841	3	23.461	5	24.928	4
4	23.85	6	27.643	4	30.513	5	28.677	7	29.993	5
5	29.803	6	34.762	1	36.278	5	35.608	2	37.604	4
6	35.139	2	43.172	5	42.186	5	41.75	3	43.241	3
7	41.812	5	50.939	2	50.973	4	48.115	6	49.937	4

Table 2. Parameters of five problem instances in case of twelve tasks in planning horizon

Task No.	Instance 1		Instance 2		Instance 3		Instance4		Instance 5	
	Appearance time [h]	Location	Appearance time [h]	Location	Appearance time [h]	Location	Appearance time [h]	Location	Appearance time [h]	Location
1	3.333	1	5.14	2	4.368	6	3.329	2	3.441	6
2	7.255	5	9.927	2	8.889	5	8.56	4	7.337	3
3	11.435	2	13.071	7	13.887	5	13.008	3	10.339	3
4	16.817	5	17.424	4	16.966	4	17.493	5	13.892	6
5	20.664	3	20.646	6	21.945	5	21.723	7	18.8	5
6	24.725	6	24.214	7	25.181	1	25.229	7	22.619	2
7	29.831	6	27.844	3	28.575	6	29.048	6	27.6	6
8	34.638	6	30.947	4	32.261	3	33.399	3	30.766	7
9	39.408	7	35.579	7	37.546	6	38.708	6	34.185	6
10	44.402	5	38.954	1	40.798	2	42.376	3	39.345	4
11	48.393	3	42.539	6	45.176	5	45.481	3	43.764	3
12	52.116	6	46.86	2	50.143	6	49.218	3	49.074	4

Results obtained through application of three mentioned approaches that comprise total waiting time of barges and schedules of handling devices are presented in Table 3. Table 4 shows relative ratios of solutions obtained by different solution approaches.

Results of the experiments for the case of seven tasks in planning horizon show that in contrast to the Insertion heuristics, which gave solutions that are in average worse than optimal solutions for nearly 17%, DDI gave results that are in average worse than optimal solutions only for 0.66%. Namely, solution generated with DDI had non optimal result in two instances, of which one has difference of roughly 1% compared to optimal result, while the difference of the second was around 2.5%. Table 3 shows that average solution of DDI is for around 17% better than Insertion heuristic, 51.5h compared to 62.1h. However differences are in much wider range from 3.14% to 29.28%.

In the case of twelve tasks problem instances it is almost the same since average solution of DDI heuristics is better than average solution obtained by Insertion's, for around 18% (165.1h compared to 201.8h), However, in this case Insertion heuristics gave slightly (0.66%) better result for second instance.

Also, from the solutions obtained by DDI heuristics, presented in Table 3, in case of fourth instance it is noticeable that solution could be improvemed if servicing sequences of the second and third device are inter changed. In that way device with higher productivity would serve more tasks, meaning that serving time would be shorter for 3.33 hours (2.46%) in that case.

5. CONCLUSION

It can be concluded that the proposed DD heuristics is very simple for application, and gives good solutions in a very short time, which suggest a possibility for practical use. Since the heuristic is promising, and gives encouraging results, it is necessary to continue with its implementation to additional examples, particularly to larger problems (for example 20 nodes), which could be more appropriate as typical dimension of the real barges' unloading problem. Another research direction is in the heuristic performances determination, and its adoption to real system requirements, particularly in the context of node priorities as well as in improving performance measures used for determining heuristic solution. Also, introducing solution improvement phase through interchanging tasks sequences, and devices would also be possible direction of future research.Additional research opportunity is a dynamic problem solving.

Finaly, overall conclusion that can be drawn from this study is that it's deserves conducting a more comprehensive study which will include larger number of instances realized on different system layouts with different distributions of tasks' inter arrival times.

ACKNOWLEDGMENT
This work was partially supported by Ministry of Science and technological Development Republic of Serbia, through the project TR 15018, for the period 2008-2010.

6. REFERENCES

Bjelic N., (2009) Operational Control Models of Material Handling Systems with Distributed Resources, *MSc Thesis*, University of Belgrade, Serbia (in Serbian)

Christiansen, M., Fagerholt, K. , Nygreen, B., Ronen, D., (2 007), Maritime tr ansportation, *Chapter 4 in Handbook of Operations Research and Management Science* 14.

van der Meer R., (2000), Operational Control of Internal Transport, *PhD Thesis*, Erasmus University, Roterdam , The Netherlands

O'Brien, G., Crane, R., (1959), The Scheduling of a Barge Line, *Operations Research* 7/5, 561-570

Schwartz, N. L., (1968), Discrete programs for moving known cargos from origins to destinations on time at minimum barge line f eet cost, *Transportation Science* 2, 134–145

Taylor, G. D., White, T.C., DePuy, G. W., Dr osos, D.J ., (2005.), A simulation-based s oftware s ystem for barge dispatching and boat assignment in inland waterways, *Simulation Modelling Practice and Theory* 13, 550-565

Vidovic M., Vukadinovic K., (2006), Allocation planning of handling devices for barges unloading, *Proceedings of the 11th Meeting of the EURO Working Group on Transportation*, Bari, 740–747

Vukadinovic, K., Teo dorovic, D ., (1994), A fuzzy a proach to vessel dispatc hing problem, *European Journal of Operational Research* 76, 155-164

Vukadinovic, K., Teodorovic, D., Pavk ovic, G., (1997), A neural network aproach to the v essel dispatching problem, *European Journal of Operational Research* 102, 473-487.

Table 3. Solutions of the testing problems obtained by mathematical model, Insertion heuristics and DDI heuristics

		Seven task planning horizon						Twelve task planning horizon			
		Optimum		Insertion		DDI		Insertion		DDI	
Instance	Device	Task assignments	Objective function	Task assignments	Objective function	Task assignments	Objective function	Task assignments	Objective function	Task assignments	Objective function
1	1	[2]		[3;6]		[3]		[3;5;10]		[1;5;11]	
	2	[3;6]	49.85	[2;5]	51.414	[2;6]	49.85	[2;4;6;12;9]	225.52	[2;4;10]	169.03
	3	[1;4;5;7]		[1;4;7]		[1;4;5;7]		[1;7;8;11]		[3;6;7;8;12;9]	
2	1	[3]		[3;7]		[5]		[3;6;9]		[3;9;6]	
	2	[2;4;6]	45	[2;5]	53.016	[3;7]	45.43	[2;5;8;12;10]	188.83	[2;8;12;10]	190.09
	3	[1;5;7]		[1;4;6]		[1;2;4;6]		[1;4;7;11]		[1;4;5;11;7]	
3	1	[1]		[3;7]		[]		[3;7;9;12]		[6;8]	
	2	[3;7]	61.667	[2;4;5]	79.243	[1;3;7]	63.16	[2;5;8]	175.37	[2;3;5;10]	139.7
	3	[2;4;5;6]		[1;6]		[2;4;5;6]		[1;4;6;10;11]		[1;4;7;9;12;11]	
4	1	[4]		[3;7]		[4]		[3;8;10;11;12]		[5;6]	
	2	[2;3;7]	55.617	[2;5;6]	70.832	[2;3;7]	55.617	[2;4;7;9]	188.93	[1;3;8;10;11;12]	135.31
	3	[1;5;6]		[1;4]		[1;5;6]		[1;5;6]		[2;4;7;9]	
5	1	[4]		[3;6]		[4]		[3;8]		[2;3;11]	
	2	[2;6]	43.333	[2;5]	56.0203	[2;6]	43.333	[2;6;10;11]	230.33	[6;10;5]	191.4
	3	[1;3;5;7]		[1;4;7]		[1;3;5;7]		[1;4;7;9;12;5]		[1;4;7;9;12;8]	
Average			51.0934		62.10506		51.478		201.796		165.106

Table 4. Relative relations between solutions

	Seven task planning horizon			Twelve task planning horizon
Relation	(Insertion-Optimal)/ Insertion	(DDI-Optimal)/ DDI	(DDI-Insertion)/ DDI	(DDI-Insertion)/ DDI
1st instance	3.04%	0.00%	-3.14%	-33.42%
2nd instance	15.12%	0.95%	-16.70%	0.66%
3rd instance	22.18%	2.36%	-25.46%	-25.53%
4th instance	21.48%	0.00%	-27.36%	-39.63%
5th instance	22.65%	0.00%	-29.28%	-20.34%
Average	16.89%	0.66%	-20.39%	-23.65%

A SYUDY ON THE PLANNING OF LINER SHIP SERVICE WITH SHIP SPEED AND BUNKERING PORT CONSIDERATIONS

Zhishuang Yao[1], Szu Hui Ng[2*], and Loo Hay Lee[3]

[1,2,3]Department of Industrial and Systems Engineering, National University of Singapore
10 Kent Ridge Crescent, 119260, Singapore
*Correspondence: isensh@nus.edu.sg

Abstract: We consider a shipping liner service planning problem which incorporates ship speed determination and bunkering port selection. We first provide an empirical formula to express the relationship between bunker fuel consumption rate and ship speed for different sizes of containerships based on real data. We then develop a planning level optimization model to minimize total bunker fuel cost for single shipping liner service considering different bunker fuel prices at different bunkering-available ports and port arrival time windows. Based on the model, we study the effects of port arrival time window, bunker fuel price and ship bunker fuel capacity on the ship speed and bunkering port decisions of a single service. We finally provide some insights obtained from a case study.

1. INTRODUCTION

Liner shipping industry has increased significantly during recent decades. One main challenge currently faced by the shipping liners is how to reduce bunker fuel cost. In recent years, the bunker fuel costs have increased to more than half of a shipping company's total operation costs. Therefore, reduction of bunker fuel cost can bring substantial savings to the total operation costs. Besides using some technical improvement strategies (e.g. using more efficient main engine), shipping liners normally adopt the slow steaming strategy to reduce the bunker fuel consumption. However, slow streaming throughout the whole trip will increase the total traveling time, which in turn may affect the schedule of the liner service. Therefore, how to adjust the ship speed at each leg of the whole service route becomes an important issue. Furthermore, bunker fuel prices at different ports can be very different even on the same day. For example, the maximum difference on the same day can be more than US$60 per metric ton (see

Table **1**). In addition, some ports provide price discounts for refueling. The problem of where to refuel is therefore another important consideration. In this paper, we consider a shipping liner service planning problem which incorporates the ship speed determination and bunkering port selection. The effects of port arrival time window, bunker fuel price and ship bunker fuel capacity on the ship speed and bunkering port decisions of a single service are also studied. Finally, some insights are provided.

Table 1. One day bunker fuel prices (380 CST) at different ports (US$/ton) (Data source: Bloomberg 2009)

Date	Port					Max difference
	Singapore	Fujairah	Hong Kong	Rotterdam	Houston	
3-Sep-2004	188	181.5	199	161.5	167	37.5
3-Sep-2005	336	319	341	303	311	38
3-Sep-2006	291	286	313	288	312	27
3-Sep-2007	379.5	397.5	389.5	352	355	45.5
3-Sep-2008	677.5	679.5	673.5	619	650	60.5
3-Sep-2009	433.5	443.5	440.5	424	422	21.5

The study on operations research (OR) based maritime transportation has attracted immense interests in recent decades. Ronen (1983) provided the first survey in ship routing and scheduling problems. Ronen (1993) and Christiansen et al. (2004) reviewed the study on the similar problem in the literature published in the decade 1982-1992 and the decade 1993-2003 respectively. According to Christiansen et al. (2007), the growing number of the references in the above three review papers shows the increasing research interest in OR-based maritime transportation. Among these studies however, there are only a few works related to ship speed and ship bunkering related issues. Ronen (1982) analyzed the tradeoff between fuel savings through slow steaming and the loss of revenues due to the resulting voyage extension. Perakis and Papadakis (1987a, 1987b) determined the fleet deployment and the associated optimal speed for ships between two ports. Fagerholt (2001) considered a multi-ship scheduling problem for speed determination with soft time windows. Christiansen and Fagerholt (2002) introduced a robust ship scheduling problem with the consideration of multiple time windows, where ship speed is an implicit decision incorporated in the ship schedule. Notteboom and Vernimmen (2009) provided a descriptive model to study the effect of increasing bunker fuel costs on the design of liner services. Fagerholt et al. (2010) studied a bunker fuel consumption minimization problem for a fixed shipping route with time windows by optimizing ship speeds at each leg of the route. Ronen (2010) analyzed the tradeoff between ship speed reduction and adding additional ships to a container liner route. To the best of authors' knowledge, there is no study determining ship speed at each leg and bunkering ports together for a single service route. While these decisions are important in the minimization of bunker fuel cost in an environment where different bunkering-available ports offer different bunker fuel prices. Therefore, we propose a shipping liner service planning model to determine ship speeds and bunkering ports in order to minimize total bunker fuel cost.

The rest of this paper is organized as follows: in the next section we describe the problem and present an empirical formula to express the relationship between bunker fuel consumption rate and ship speed. Following this, we present the mathematical model and provide a case study and some insights from the study. Finally, we conclude this paper.

1.1 DESCRIPTION OF THE PROBLEM

We consider a single shipping liner service, where the service route, service frequency, port times and number of ships deployed in the service are known. We assume that all ships deployed in the service are homogenous as this is not uncommon in practice. Therefore, the total cycle time for a ship in the service is known (e.g. if N ships are deployed in a weekly service, the total cycle time for a ship is N*7*24 hours). Although liner service incurs different types of costs, e.g., capital cost, operating cost (except bunker fuel cost), container cost and administration cost, these costs are constants in our study as both the number of ships deployed in the service and the total cycle time for a ship are constants. We therefore only consider bunker fuel cost. In order to reduce the bunker fuel cost, we need to adjust ship speed at each leg of the route while keeping the total cycle time unchanged. Furthermore, bunker fuel prices at different ports are different and some ports provide price discounts under certain contractual agreements, hence, we also need to decide where to bunker the ship. Therefore, we consider a shipping liner service planning problem which incorporates the ship speeds determination and bunkering ports selection. The effects of port arrival time window, bunker fuel price and ship bunker fuel capacity on the ship speeds and bunkering ports of a single service are also studied.

In order to calculate the bunker fuel consumption of a ship, we first need to study the relationship between bunker fuel consumption rate and ship speed. Although some previous studies use third power of the ship speed (Ronen, 1982, Ronen, 2010) or an empirical figure (Notteboom and Vernimmen, 2009) or quadratic convex function (Fagerholt et al., 2010) to approximate the bunker fuel consumption, there is no study providing empirical formula for the relationship between bunker fuel consumption rate and ship speed for different sizes of containerships. Based on the data we acquired from a shipping liner, we analyzed this relationship for different sizes of containerships. We find that $F = k_1*V^3 + k_2$ is a good empirical formula, where F denotes the bunker fuel consumption rate (tons/day), V denotes the ship speed (knots) and k_1, k_2 are coefficients. The coefficients k_1 and k_2 for different sizes of containerships are shown in Table 2. From Table 2, we find that k_1, k_2 are different for different sizes of containerships and the general trend for k_1, k_2 is increasing with sizes of containership, which indicates that the bunker fuel consumption increases much faster with speed for the larger ships. Note also from Table 2 that the empirical formulas are valid for certain speed intervals. Compared to the existing studies, our empirical formula has the following two main differences. First, our formula has the coefficient k_2 which can make the approximation more accurate. Second, we find that the coefficients of the formula are different for different sizes of containerships, while was ignored by most existing studies.

Table 2. Coefficients for different sizes of containerships

Size (TEU)	k_1	k_2	Amount of data	Speed interval of the data (knots)
0-1000	0.004476	6.17	73	(10.5, 16.5)
1000-2000	0.004595	16.42	65	(12.5, 19.5)
2000-3000	0.004501	29.28	51	(13.5, 21)
3000-4000	0.006754	37.23	82	(14.5, 21.5)
4000-5000	0.006732	55.84	193	(15, 24)
5000-6000	0.007297	71.4	170	(14, 24)
6000-7000	0.007473	77.77	13	(18, 25)

2. MODEL DEVELOPMENT

In this section, we first provide the notations used, then present the model followed by the solution method.

Parameters

n number of port of calls;

$d_{i,j}$ distance between port i and port j (nautical miles);

t total cycle time (hours);

t_i port time at port i (hours);

e_i earliest arrival time at port *i* (hour);
l_i latest arrival time at port *i* (hour);
a_i bunker fuel consumption of a single ship at port *i* (tons);
w bunker fuel capacity of a single ship (tons);
v_{min} minimum ship speed (knots: nautical miles/hour);
v_{max} maximum ship speed (knots);
k_1, k_2 bunker fuel consumption coefficients;
q maximum number of bunkering times;
p_i^1 unit price of bunker fuel at port *i* (\$/ton);
p_i^2 , p_i^3 discounted unit prices of bunker fuel at port *i*, $p_i^3 \le p_i^2 \le p_i^1$ (\$ /ton);

Decision variables
$V_{i,j}$ ship speed between port *i* and port *j* (knots);
S_i bunker fuel order-up-to-level for a single ship at port *i* (tons);
I_i bunker fuel inventory when arriving at port *i* (tons);
B_i = 1, bunkering at port *i*; = 0, otherwise.
$\alpha_i^1,...,\alpha_i^4$ variables used for piecewise bunkering cost function at port *i*;
$\beta_i^1,...,\beta_i^3$ variables used for piecewise bunkering cost function at port *i*.
Dependent variables
C_i bunkering cost at port *i* (\$);
$F_{i,j}$ bunker fuel consumption rate between port *i* and port *j* (tons/day);
A_i arrival time at port *i* (hour).

Note that we consider incremental quantity discounts on the bunker fuel price, which is expressed by

$$C_i = \begin{cases} p_i^1 \cdot (S_i - I_i) & 0 \le S_i - I_i \le 1000 \\ p_i^1 \cdot 1000 + p_i^2 \cdot (S_i - I_i - 1000) & 1000 < S_i - I_i \le 2000 \\ p_i^1 \cdot 1000 + p_i^2 \cdot 1000 + p_i^3 \cdot (S_i - I_i - 2000) & 2000 < S_i - I_i \end{cases}$$

The planning model for single liner service is presented as follows:

$$\text{Min} \sum_{i=1}^{n} C_i$$

Subject to

$$C_i = \alpha_i^1 \cdot 0 + \alpha_i^2 \cdot 1000 \cdot p_i^1 + \alpha_i^3 \cdot (1000 \cdot p_i^1 + 1000 \cdot p_i^2) + \alpha_i^4 \cdot \left(1000 \cdot p_i^1 + 1000 \cdot p_i^2 + (w - 2000) \cdot p_i^3\right) \quad \text{for all } i \quad (1)$$

$$S_i - I_i = \alpha_i^1 \cdot 0 + \alpha_i^2 \cdot 1000 + \alpha_i^3 \cdot 2000 + \alpha_i^4 \cdot w \quad \text{for all } i \quad (2)$$

$$\alpha_i^1 \le \beta_i^1 \quad \text{for all } i \quad (3)$$

$$\alpha_i^j \le \beta_i^{j-1} + \beta_i^j \quad j = 2, 3; \text{ for all } i \quad (4)$$

$$\alpha_i^4 \le \beta_i^3 \quad \text{for all } i \quad (5)$$

$$\alpha_i^1 + \alpha_i^2 + \alpha_i^3 + \alpha_i^4 = 1 \quad \text{for all } i \quad (6)$$

$$\beta_i^1 + \beta_i^2 + \beta_i^3 = 1 \qquad \text{for all } i \qquad (7)$$

$$\alpha_i^j \geq 0 \qquad j = 1, 2, 3, 4; \text{ for all } i \qquad (8)$$

$$\beta_i^j = 0 \text{ or } 1 \qquad j = 1, 2, 3; \text{ for all } i \qquad (9)$$

$$I_1 = S_n - a_n - F_{n,n+1} \cdot d_{n,n+1} / (24 \cdot V_{n,n+1}) \qquad (10)$$

$$I_i = S_{i-1} - a_{i-1} - F_{i-1,i} \cdot d_{i-1,i} / (24 \cdot V_{i-1,i}) \qquad i = 2, \ldots, n \quad (11)$$

$$S_i - I_i \geq B_i \cdot 20\% \cdot w \qquad \text{for all } i \quad (12)$$

$$S_i - I_i \leq B_i \cdot w \qquad \text{for all } i \quad (13)$$

$$\sum_{i=1}^{n} B_i \leq q \qquad (14)$$

$$I_i \geq 5\% \cdot w \qquad \text{for all } i \quad (15)$$

$$S_i \leq w \qquad \text{for all } i \quad (16)$$

$$F_{i,i+1} = k_1 \cdot V_{i,i+1}^3 + k_2 \qquad \text{for all } i \quad (17)$$

$$v_{\min} \leq V_{i,i+1} \leq v_{max} \qquad \text{for all } i \quad (18)$$

$$A_i + t_i + d_{i,i+1} / V_{i,i+1} = A_{i+1} \qquad \text{for all } i \quad (19)$$

$$e_i \leq A_i \leq l_i \qquad \text{for all } i \quad (20)$$

$$A_1 = 0 \qquad (21)$$

$$A_{n+1} = t \qquad (22)$$

$$B_i = 0 \text{ or } 1 \qquad \text{for all } i \quad (23)$$

where $F_{n,n+1} = F_{n,1}$, $d_{n,n+1} = d_{n,1}$, $V_{n,n+1} = V_{n,1}$.

The objective function is to minimize the total bunkering costs incurred by one ship per cycle, where the bunkering cost is expressed by the standard piecewise linearization constraints (1) – (9). Constraint (10) and (11) are flow conservation constraints. Constraint (12) ensures that the minimum bunkering amount is a certain percentage of the bunker fuel capacity of a ship. Constraint (13) ensures that the bunkering amount does not exceed the bunker fuel capacity of a ship. Constraint (14) restricts the maximum number of bunkering times. Constraint (15) ensures that bunker fuel inventory has a certain minimum level. Constraint (16) ensures that the bunker fuel capacity of a ship is not violated. Constraint (17) reflects the relationship between bunker fuel consumption rate and ship speed. Constraint (18) ensures that the ship speed is within a certain interval. Constraint (19) is a time flow constraint. Constraint (20) ensures that ship arrives at a port during the time window. Constraints (21) and (22) impose the ship arrival times at the starting port. Constraint (23) is a binary constraint. Note that this model is applicable for all ships deployed in the service.

The above model is solved as follows: Defining $W_{i,i+1}=1/V_{i,i+1}$ for all i, replace $V_{i,i+1}$ in the above model by $1/W_{i,i+1}$. Then the only nonlinearity in the model is $1/W_{i,i+1}^2$, which can be approximated by a piecewise linear function of $W_{i,i+1}$. We then use a commercial solver like ILOG/CPLEX to solve the mixed integer linear programming model.

3. CASE STUDY

We now apply our model and solution approach to a service route Asia-Europe Express (AEX), which is provided by a shipping company. Figure 1 shows the route of the AEX service. Table 3 provides relevant parameters for the AEX service. We study the effects of port arrival time windows, bunker fuel prices and ship bunker fuel capacity on the ship speed, and bunkering port decisions of the service. The model and the solution procedure are coded in C ++ programming. The coded C++ program calls the ILOG/CPLEX11.0 solver to solve the MIP model.
\

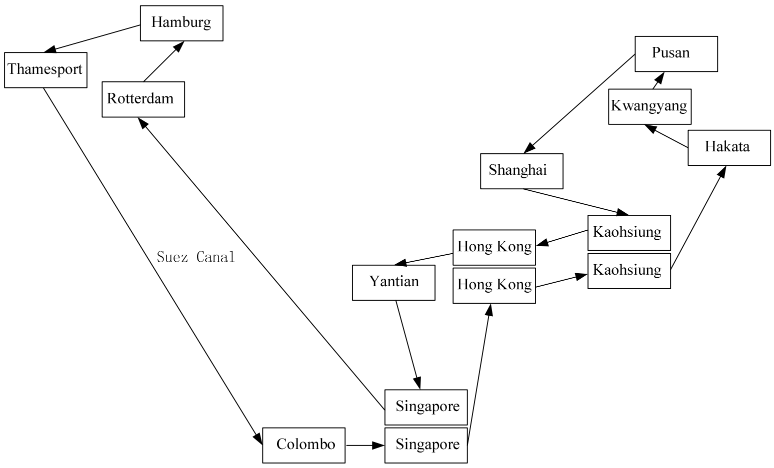


Figure 1. AEX service route

Table 3. Parameters for AEX service

Parameter	Value
Number of port of calls	15
Service frequency	weekly
Number of ships deployed	9
Ship size	6000 TEU
Total cycle time	63*24 = 1512 h
Ship bunker fuel capacity	3000 tons
Ship speed interval	(14, 24) knots
Bunker fuel consumption formula[a]	F = 0.007297 V^3 + 71.4
Maximum bunkering times	5
Bunker fuel price discount factor[b]	10% 20%

[a]: Empirical formula for 5000-6000 TEU is chosen as 6000 TEU is closer to the size of the ships used for 5000-6000 TEU fitting compared to 6000-7000 TEU, [b]: $p_i^2 = p_i^1 \cdot (1-10\%)$, $p_i^3 = p_i^1 \cdot (1-20\%)$

We first study the effects of port arrival time windows. Under current bunker fuel price settings at all bunkering-available ports and a fixed total cycle time, we compare two situations: one that considers a given port arrival time window and another that relaxes the port arrival time windows. We assume that a penalty is not incurred if we relax the port arrival time windows. We find that the optimal total bunkering cost decreases from $ 2,737,040 to $2,705,660 (-$31,380, -1.15%) if we relax the port arrival time windows. Table 4 shows the ship speeds and bunkering ports of the above two situations. In the current situation, the ship speed of the westbound voyage is relatively high because the shipping company wants to provide a fast transit time from Southeast Asia to Rotterdam. In order to reduce the bunker fuel cost, we observe from Table 4 that it is optimal to decrease the speed of the westbound voyage and increase the speed of the eastbound voyage after relaxing the port arrival time widows. The results also show that the bunkering ports change from Yantian, Rotterdam and Colombo to Kaohsiung, Rotterdam and Colombo. This change is because the increase in speed of the eastbound voyage increases the bunker fuel consumption of that voyage, requiring the ship to refuel earlier, i.e., bunkering at Kaohsiung instead of Yantian.

Table 4. AEX ship speed and bunkering ports comparisons (with/without port arrival time windows)

Port	With port arrival time window		Without port arrival time window	
	Ship speed (knots)	Bunkering amount (tons)	Ship speed (knots)	Bunkering amount (tons)
(1) Hakata	19.04		**20.82**	
(2) Kwangyang	19.03		**20.82**	
(3) Pusan	**21.69**	2	0.82	
(4) Shanghai	**21.69**	2	0.82	
(5) Kaohsiung	**21.68**	2	0.99	2850.0
(6) Hong Kong	**22.07**	2	1.02	
(7) Yantian	**23.09**	2810.1 2	1.02	
(8) Singapore	**22.09**	2	1.00	
(9) Rotterdam	**21.19**	2850.0 2	0.82	2850.0
(10) Hamburg	20.31		**20.82**	
(11) Thamesport	19.78		**20.81**	
(12) Colombo	19.79	877.9	**20.82**	806.7
(13) Singapore	19.79		**20.82**	
(14) Hong Kong	19.76		**20.82**	
(15) Kaohsiung	19.79		**20.82**	

We next study the effects of bunker fuel price. We consider six different situations of the bunker fuel price evolution along the bunkering-available ports: (1) current bunker fuel prices for all bunkering-available ports; (2) same price for all bunkering-available ports; (3) increasing prices along the service route; (4) decreasing prices along the service route; (5) increasing then decreasing prices and (6) decreasing then increasing prices. For a fair comparison, we set the average prices of all bunkering-available ports to be the same (US$460/ton) in all situations. Table 5 shows the results of the optimal bunkering ports and amounts for the above six situations. From the results, we find that there are always only three bunkering ports. This is mainly due to the ship bunker fuel capacity (3 000 tons) used. We also observe from Table 5 that Rotterdam is always a bunkering port. This can be explained by the fact that ships go through a very long voyage from Singapore to Rotterdam and would have consumed a great amount of bunker fuel by the time they arrive at the Rotterdam port, thus requiring the ships to be refueled at the Rotterdam port. Moreover, we see that the port with the lowest bunker fuel price may not be the optimal bunkering port, as shown in situations (3), (4) and (5). This is mainly due to the different bunker fuel requirements at different legs of the service route.

Table 5. Optimal bunkering ports and amounts (AEX) for different situations

Port	Bunkering amount (tons)					
	Situation (1)	Situation (2)	Situation (3)	Situation (4)	Situation (5)	Situation (6)
(6) Hong Kong			2850.0		846.8	
(7) Yantian	2810.1					2810.0
(8) Singapore	84	6.8	84	6.8		
(9) Rotterdam	2850.0 2	850.0	2828.1 2	841.3 2	841.3	2850.0
(10) Hamburg						
(11) Thamesport						
(12) Colombo	877.9 85		9.9		87	7.9
(13) Singapore						
(14) Hong Kong		2841.3		2850.0	2850.0	
(15) Kaohsiung						

Following, we study the effects of ship bunker fuel capacity. Under a given bunker fuel price settings and keeping the total cycle ti me unchanged, we find that the optimal total bunkering cost decreases from $2,737,040 to $2 ,616,920 (-$120,120, -4.39%) if we can increase the ship bunk er fuel cap acity from 3 000 tons to 4000 ton s. Results also show that th e number of bunkering po rts d ecreases from three (Y antian, Rotterdam and C olombo) t o tw o (Pusan an d Rotterdam). This is mainly due to the bunker fuel price discounts offered. From all results, we can draw some insights. First, the evolution of the bunker fuel prices at bunkering-available ports is important to the selection of bunkering ports. Different ev olution of bunker fuel prices may lead to different optimal bunkering ports, which shows that t he current practice (which usually sees s hips bunkering at o nly a f ew fixed p orts) ca n b e im proved. S econd, by c hanging the arrival times at ports and the speeds at legs while keeping the total cy cle time unchanged, the bunker fuel cost ca n be further reduced from that of the optimal schedule under the current given port arrival time windows. Third, increasing the shi p bunke r fuel capacity can bring re latively significa nt be nefits for ship ping liners in term s of bunker fu el co st savings. This is a useful consideration in the deployment of ships or in the design of new ships.

4. CONCLUSIONS

In this paper, we study a lin er service planning problem which considers ship speed determination and bunkering port selection. We first provide an empirical formula to expres s the relationship between bunker fuel consumption rate and ship speed. We then develop a single se rvice planning le vel optimization model to determine bunkering ports and ship speeds. Based on the model, we study the effects of port arrival time windows, bunker fuel prices and ship bunker fuel capacity on th e ship s peed and bunkering port decisions of a si ngle service. We finally provide some insights. Future research can incorporate some uncertainties into the model or consider an operational level model.

ACKNOWLEDGEMENTS

This research was carried out with financial support from Neptune Orient Lines (NOL) fellowship. This support is gratefully acknowledged.

5. REFERENCES

Christiansen, M. and Fagerholt, K. (2002). Robust Ship Scheduling with Multiple Time Windows. *Naval Research Logistics, 49:* 611–625.

Christiansen, M., Fagerholt, K., Nygreen, B. and Ronen, D. (2007). Maritime Transportation. In: Barnhart, C. and Laporte, G. (eds). *Transportation. Handbooks in Operations Research and Management Science*, Vol. 14. North-Holland, Amsterdam, 189–284.

Christiansen, M., Fagerholt, K. and Ronen, D. (2004). Ship Routing and Scheduling: Status and Perspectives. *Transportation Science 38(1):* 1–18.

Fagerholt, K. (2001). Ship Scheduling with Soft Time Windows: An Optimisation based Approach. *European Journal of Operational Research, 131:* 559–571.

Fagerholt, K., Laporte, G. and Norstad, I. (2010). Reducing Fuel Emissions by Optimizing Speed on Shipping Routes. *Journal of Operational Research Society, 61(3):* 523–529.

Notteboom, T. E. and Vernimmen, B. (2009). The Effect of High Fuel Costs on Liner Service Configuration in Container Shipping. *Journal of Transport Geography, 17:* 325–337.

Perakis, A. N. and Papadakis, N. A. (1987). Fleet Deployment Optimization Models. Part 1. *Maritime Policy and Management, 14(2):* 127–144.

Perakis, A. N. and Papadakis, N. A. (1987). Fleet Deployment Optimization Models. Part 2. *Maritime Policy and Management, 14(2):* 145–155.

Ronen, D. (1982). The Effect of Oil Price on the Optimal Speed of Ships. *Journal of Operational Research Society, 33(11):* 1035–1040.

Ronen, D. (1983). Cargo Ships Routing and Scheduling: Survey of Models and Problems. *European Journal of Operational Research, 12:* 119–126.

Ronen, D. (1993). Ship Scheduling: the Last Decade. *European Journal of Operational Research, 71:* 325–333.

Ronen, D. (2010). The Effect of Oil Price on Containership Speed and Fleet Size. *Journal of Operational Research Society,* (advance online publication, 13 January 2010) doi:10.1057/jors.2009.169.

Session C1: SCM 3

·Day1: Sep. 15 (Wed.)

·Time: 15:00 - 16:20

·Chair: Bongju Jeong

·Room: Camellia, 5F

GLOBAL SUPPLY CHAIN SUSTAINABILITY IN THE CONSUMER GOODS INDUSTRY

Kannegiesser, M. [1]**, Günther, H.-O.** [1*]**, Otten, D.** [1]

[1]Department of Production Management
Technical University of Berlin
Strasse des 17. Juni 135
10623 Berlin, Germany
e-mail: hans-otto.guenther@tu-berlin.de

1. INTRODUCTION

Sustainability – the balance of economic, social and environmental performance – is a key task in the 21st century given the challenges of climate change or natural resource scarcity to name a few. Supply chain network design traditionally focused on costs and profit optimization. The increasing importance of environmental objectives such as reducing the carbon footprint to achieve not only economic but overall sustainability requires enhancing supply network design approaches. We consider a global industry supply chain network for packaged food from raw material to the end consumer including reverse flows and investigate and propose the *minimize-time-to-sustainability* approach as new sustainability optimization concept to be compared with trade-off optimization concepts.

2. PROBLEM SPECIFICATION

We focus on a simplified end-to-end packaged food supply chain model (see figure 1). The supply chain comprises multiple steps from raw materials to end-consumer as well as closed-loop and recycling flows. Our focus is on two value adding activities: *manufacturing and logistics* at locations as well as *transportation* using different modes between locations. In this model-based approach, food raw materials are produced by farmers in a first step and then processed by food suppliers. Packaging raw materials such as plastics, pulp and paper are produced by process industry companies and then processed by packaging part suppliers. The consumer products manufacturer combines different processed food products to a final, packaged food product using respective packaging. Packaged food is then shipped to retailers distributing it to end-consumers via different channels such as grocery stores, food services or using e-commerce. We also consider recycling flows, where food waste is disposed and packaging waste can be disposed or recycled and sent back in a closed loop.

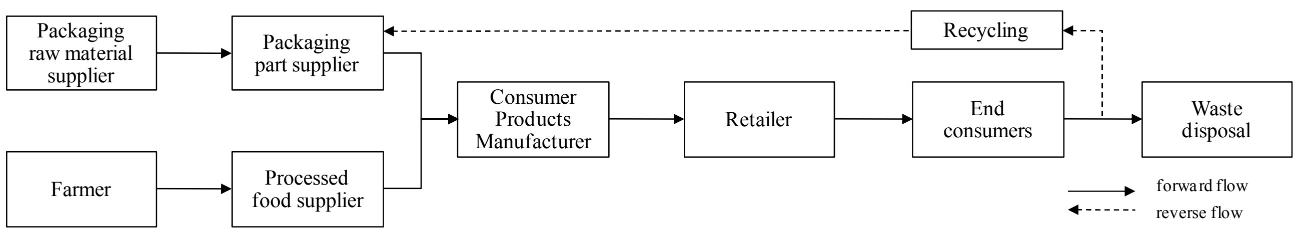


Figure 1: Supply Chain Structure

Companies in this value chain face network design decisions such as where to produce which product at which volume, how to transport and how to distribute. These decisions are related to strategic investment decisions with planning horizons of multiple years and are made on an aggregated level. While previously cost, quality and customer demand have been key parameters for decision making, sustainability criteria such as greenhouse gas emissions, waste or water consumption to name a few become now part of the optimization problem, when companies pledge voluntarily or are required by regulation to meet various and partly conflicting economic, social and environmental objectives.

We focus on the optimization of three sustainability objectives: to minimize total landed costs, CO2 emissions and total lead time in the considered industry supply chain. Influencing factors can be different by location (e.g. different labor costs in emerging countries vs. Western countries, different energy mixes leading to different CO2 emissions caused by energy consumption to name a few) and by transportation lane (e.g. different costs, CO2 emissions and lead times e.g. depending on transportation modes such as air, truck, ship or rail). Two areas influence the network design and overall sustainability objectives: *internal practices* e.g. application of best practices in the supply chain in manufacturing, logistics and transportation and *external factors* such as CO2 prices and taxes, energy prices, labor costs, efficiency increase or changing consumer behavior.

3. LITERATURE REVIEW

The problem is linked to *traditional supply chain network design*, *sustainable supply chain* in general and *industry-specific supply chain problems* for food and consumer products.

Traditional supply network design problems focus on economic optimization criteria such as maximizing profit or minimizing costs for a long-term production and distribution network design (Goetschalckx and Fleischmann, 2004). These problems have increasingly to deal with global scope and specifics such as optimizing costs and lead time differences for global production and distribution networks (c.f. Vidal and Goetschalckx, 2001) and in various industries e.g. Electronics (e.g. Arntzen *et al.*, 1995) or Chemicals (e.g. Bassett and Gardner, 2010) to name a few. Traditional network design approaches have to be enhanced now by incorporating additional sustainability criteria.

Research on *green* and *sustainable supply chain* intensified recently specifically in the context of environmental and carbon trading that many industries are confronted with. Srivastava (2007) provides a classification of green supply chain management state-of-the-art literature clustered into *green design of products* using lifecycle analysis (LCA) or environmental conscious design (ECD) as well as *green operations* with the subareas *green manufacturing and remanufacturing*, *reverse logistics and network design* as well as *waste management* either using empirical studies or mathematical modeling. Linton *et al.* (2007) provide an issue overview that forms part of sustainable supply chain management and that extends "beyond the core of supply chain management" specifically incorporating product usage phase, recycling and reverse flows at end-of-product-life. Many studies focus on structural, qualitative and analytical aspects e.g. how to design general sustainable supply chain structures, argumentation structures why a sustainable supply chain is important or analytical evaluation of a supply chain's environmental performance e.g. using product life cycle analysis (Tsoulfas *et al.*, 2002) or empirically analyze sustainable supply chain practices at companies (Zhu *et al.*, 2008). Rosič *et al.* (2009) identify and evaluate network design and transportation best practices that have an impact and can help to improve sustainability in supply chain networks such as a network redesign with optimized offshoring and nearshoring, a flexible supply base and flexible transportation with optimized modes to name a few. Venkat and Wakeland (2006) show that a cost and time-minimized lean supply chain is not necessarily green, specifically if a high frequency of deliveries is required to be lean as in the case of just-in-time delivery leading to transport emissions in transportation. Several authors started to propose general mathematical optimization models: Sheu *et al.* (2005) present a multi-objective optimization approach to balance forward vs. reverse supply chain profit focusing on costs. Quariguasi Frota Neto *et al.* (2009) propose an algorithm approach to find the efficient frontier between the trade-off of profitability and environmental criteria such as waste or CO2 emissions in general and applied the approach for a case in the pulp and paper industry (Quariguasi Frota Neto *et al.*, 2008). Ramudhin *et al.* (2008) and Diabat and Simchi-Levi (2009) specifically include carbon markets and CO2 caps and price impact on trade-off with costs when designing supply networks and optimizing modes. Summarizing, mathematical model approaches developed so far focus on optimizing the trade-offs between multiple, partly conflicting economic and environmental goals as well as trying to incorporate external factors such as CO2 markets and prices and their impact on results.

Industry-specific supply chain problems for consumer and food industry are linked to specifics such as planning of shelf life and product quality (Luetke Entrup *et al.* 2005, van Sonsbeck *et.al.* 1996) specifically in the area of fresh food production. On the distribution and retailer side, optimizing retail distribution network structures and logistic performance (Hsiao *et al.*, 2010) or transportation planning with bundling of transports and optimizing transport routes and schedules (Günther and Seiler, 2010) are in focus to name a few. Recently, sustainability criteria have been increasingly integrated into food supply chain management problems. Hamprecht *et al.* (2005) present the integration of sustainability criteria into supply chain decision making and controlling for the case of Nestlé and Wognum *et al.* (2010) discuss technical and organizational solutions to improve sustainability in the food chain.

Only little research on supply chain network design in the food sector has been conducted so far. We propose a new sustainability optimization approach that can be compared to trade-off optimization approaches developed so far.

4. SUSTAINABILITY OPTIMIZATION MODEL

The proposed sustainability optimization model is structured along manufacturing and logistics *processes* and *transportation* steps as well as *two optimization approaches*. Basic indices are periods, products, locations, processes and transportation modes. The time horizon is considered as long-term, in our case starting at 2010 as baseline year towards 2030 as final period. Products are input and output products such as farming products, processed food, packaging material, packaged organic food, food waste or packaging waste to name a few. Locations are manufacturing and logistics locations determined by the role in the industry supply chain e.g. retailer or consumer product manufacturer as well as the geographic location on an aggregated level such as import and domestic regions.

Processes are linked to a location and reflect manufacturing and/or logistics processes such as farming, food processing and manufacturing, retail warehousing and retailing in the store as forward processes as well as reverse processes such as recycling and waste disposal or composting. A process has one or multiple input and/or output products, e.g. the packaged food process of a consumer products manufacturer requires multiple input products such as processed food and packaging to produce one output product of packaged food. Processes are determined by location-specific control parameters such as minimum and maximum process volume, operating, opening and closing costs, volume-dependent emission factors and lead time factors. Since we focus on a long-term industry supply chain network design problem, the problem is not linked to one company with several locations, but to multiple companies with respective capacities for a certain value-adding process. The model decides process and product volumes, total, opened and closed units, and all costs, emissions and lead time.

Transportation is defined by product, start and end location as well as transportation mode such as rail, ship, truck or air cargo. Transportation has the same control parameter as processes, minimum and maximum volumes, cost and emission factors. We also consider transport units and costs of investing or shutting-down additional transport units and equipment. Since we focus on an end-to-end industry perspective, transport service providers and their investments, opening and closing costs are integral part of the supply chain design problem to be considered. Decision variables are transport volumes, opened and closed transport units in each period as well as respective costs, emissions and lead times. Material flows are balanced at all transport start and end locations with the input and output product volumes at the respective locations.

Two optimization approaches are presented: a *multi-objective trade-off* approach and a *minimize-time-to-sustainability* approach. The multi-objective trade-off approach minimizes costs, emissions and lead time all weighted with respective weighting factors. Weighting factors are varied to reflect different importance of the individual criteria of the objective function and to generate trade-off curves. Alternatively we propose the concept of *minimize-time-to-sustainability.* Time-to-sustainabilty is the time required when *all* sustainability criteria have reached a defined target level *sustainably.* A maximum sustainable level for costs, emissions and lead time has to be defined, e.g. if 40% carbon emission reduction until the planning horizon of 2030 is the required sustainable level e.g. to limit the temperature increase to 2° degree Celsius. The model would try to find the best network design development paths to achieve the 40% target as early as possible *together* with sustainable costs and lead time levels. Important is that a period can only be *sustainable* if all subsequent periods are also sustainable not returning back to unsustainable levels. The time-to-sustainability concept allows analyzing feasibility of sustainability targets by varying different target levels and comparing sustainability time results. Both strategies are evaluated in numerical evaluations.

5. NUMERICAL EVALUATION FOR PACKAGED FOOD CASE

We evaluate the model and the two discussed optimization strategies with realistic data for an end-to-end packaged food industry supply chain from farming to end-consumer usage including recycling and waste disposal. Case data are generated based on public statistics, company information and reports as well as reference cases covering the characteristics described in section 2.

We implemented the model using *IBM ILOG OPL Studio 6.3®*, *IBM ILOG CPLEX 12.1®* and *Microsoft Access 2003®* on an *Intel® Xeon®* processor with 2.49 GHz and 4 GB RAM. Solving times are efficient between 1 to max. 5 seconds for 10 periods of data.

We conducted baseline, multi-objective trade-off and minimize-time-to-sustainability evaluations. The baseline evaluation is used to generate a reference scenario; other optimization strategies can be compared against it. In the multi-objective trade-off evaluation, we apply different weights for costs, emissions and lead time and generate trade-off curves between the three objectives. The evaluations clearly show the trade-offs between costs and emissions and the respective optimum network designs. Thirdly, we apply the minimize-time-to-sustainability approach and vary different sustainability targets for costs, emissions and lead-time e.g. from 100% - meaning the baseline level is already sustainable – down to 20% meaning, a sustainable level e.g. of CO2 emissions is 20% of the baseline emission representing an 80% emission reduction.

We show that multi-objective trade-off optimization is less appropriate for long-term network design decision making: companies rather set long-term sustainability targets e.g. CO2 reduction targets and need to know if, how and when these targets together with other targets can be achieved. Here, the integral meaning of sustainability comes into play, which is about finding the balance in a system, where *all* parameters have achieved a sustainable state that the system can be handed over to next generations without limiting their needs and capabilities in the future. The time-to-sustainability approach shows the earliest period, when a sustainable steady state for costs, CO2 and leadtime is achieved. Therefore, the results are more meaningful for a long-term network design and show the feasibility of sustainability targets as well as the sensitivity of network design results confronted with different sustainability target levels.

6. SUMMARY AND CONCLUSIONS

We propose the *minimize-time-to-sustainability* approach as a new concept to integrate sustainability measures and targets into long-term supply chain network design. The approach is applied for optimizing costs, emissions and lead time in manufacturing, logistics and transportation. It is compared with most commonly used multi-objective optimization approaches balancing trade-offs between multiple sustainability goals. Applied to a packaged food industry supply chain, the minimize-time-to-sustainability approach allows to determine the feasibility of long-term sustainability targets and to identify the optimal supply chain design path to achieve targets in shortest time. It can be easily expended by integrating additional – originally incompatible - sustainability criteria such as waste, water consumption or social objectives. It can be expanded to further industry case studies as well as integrating scenario-based planning for uncertain, external factors such as CO2 taxes or energy prices.

REFERENCES

Arntzen, B., Brown, G., Harrison, T., Trafton, L. (1995). Global Supply Chain Management at Digital Equipment Corporation. *Interfaces, 25 (1),* 69-93.

Bassett, M., Gardner, L. (2010). Optimizing the design of global supply chains at Dow AgroSciences. *Computers and Chemical Engineering, 34 (2):* 254-265.

Diabat, D., Simchi-Levi, D. (2009). A Carbon-Capped Supply Chain Network Problem. *Proceedings of the 2009 IEEM Industrial Engineering and Engineering Management,* 523-527.

Goetschalckx, M., Fleischmann, B. (2008) Strategic Network Planning. In: H. Stadtler, C. Kilger (Ed.), Supply Chain Management and Advanced Planning, 4th edition, *Springer, Berlin et al.,* 117-132.

Günther, H.-O., Seiler, T. (2010). Operative Transportation Planning in Consumer Goods Supply Chains. *Flexible Services and Manufacturing Journal, 21 (1-2):* 51-74.

Hamprecht, J., Corsten, D., Noll, M., Meier, E. (2005). Controlling the sustainability of food supply chains. *Supply Chain Management, 10 (1):* 7-10.

Hsiao, H., Kemp, R., van der Vorst, J., Omta, S. (2010). A classification of logistic outsourcing levels and their impact on service performance: Evidence from the food processing industry. *International Journal of Production Economics*, *124:* 75-86.

Linton, J., Klassen, R., Jayaraman, V. (2007). Sustainable supply chains: An introduction. *Journal of Operations Management, 25 (6):* 1075-1082.

Lütke Entrup, M., Günther, H.-O., van Beek, P., Grunow, M., Seiler, T. (2005), MILP approaches to shelf life integrated planning and scheduling in yogurt production. *International Journal of Production Research, 43:* 5071-5100.

Quariguasi Frota Neto, J., Bloemhof-Ruwaard, J., van Nunen, J., van Heck, E. (2008). Designing and evaluating sustainable logistics networks. *International Journal of Production Economics, 48 (15):* 4463-4481.

Quariguasi Frota Neto, J., Walther, G., Bloemhof, J., van Nunen, J., Spengler, T. (2009). A methodology for assessing eco-efficiency in logistics networks. *European Journal of Operational Research, 193 (3):* 670-682.

Ramudhin, A., Chaabane, A., Kharoune, M., Paquet, M. (2008). Carbon Market Sensitive Green Supply Chain Network Design. *IEEE International Conference on Industrial Engineering and Engineering Management,* 1093-1097.

Rosič, H., Bauer, G., Jammernegg, W. (2009). A framework for economic and environmental sustainability and resilience of supply chains. In: Reiner G., Rapid Modeling for Increasing Competitiveness. *Springer, London.* 91-104.

Sheu, J., Chou, Y., Hu, C. (2005). An integrated logistics operational model for green-supply chain management. *Transportation Research Part E: Logistics and Transportation Review, 41 (4):* 287-313.

Srivastava, S. (2007). Green supply-chain management: A state-of-the-art literature review. *International Journal of Management Reviews, 9 (1):* 53-80.

Tsoulfas, G., Pappis, C., Minner, S. (2002). An environmental analysis of the reverse supply chain of SLI batteries. *Resources Conservation and Recycling, 36 (2):* 135-154.

Wognum, P., Bremmers, H., Trienekens, J., van der Vorst J., Bloemhof J. (2010). Systems for sustainability and transparency of food supply chains – Current status and challenges. *Advanced Engineering Informatics, doi:10.1016/j.aei.2010.06.001 (in press):* 1-12.

van Sonsbeek, J., van Beek, P., Urlings.H., Bijker, P., Hagelaar, J. (1996). Mixed integer programming for strategic decision support in the slaughter by-product chain. *OR Spektrum, 19:* 159-168.

Venkat K., Wakeland W. (2006). Is Lean Necessarily Green? *Proceedings of the 50th Annual Meeting of the ISSS.* http://journals.isss.org/index.php/proceedings50th/article/view/284/67.

Zhu, Q., Sarkis, J., Lai, K.-H. (2008). Confirmation of a measurement model for green supply chain management practices implementation. *International Journal of Production Economics, 111 (2):* 261-273.

GREENHOUSE GAS EMISSIONS FOR THE TRANSPORTATION SECTOR IN KOREA

Seungbae Sim[1], Jisoo Oh[1], and Bongju Jeong[1*]

[1]Department of Information and Industrial Engineering
Yonsei University
134 Shinchon-dong, Seodaemun-gu
Seoul, Korea
e-mail: bongju@yonsei.ac.kr

Abstract: The objective of this paper is to develop a methodology of calculating the greenhouse gas emissions for transportation sector and apply the proposed methodology in Korea. Transportation sector is one of the most critical sectors in view of environmental pollutions and represents about 20 percent of yearly total greenhouse gas emissions for energy sector in Korea. In spite of the importance, the methodology to analyze greenhouse gas emissions uses a Tier 1 method which is the simplest one among the methodologies recommended by IPCC guideline. We define and develop a methodology of calculating the greenhouse gas emissions for the transportation sector considering the key factors, transportation infrastructure and operational environments in Korea. The proposed methodology can analyze the amount of greenhouse gas emissions for each transport mode and integrate it through transportation environments. It also provides the decision-making tool for government to make the transportation polices. As a result, Korean government can effectively manage the transportation sector and improve national eco-environment.

1. INTRODUCTION

1.1 Global issues on environmental impact

In a recent decade, there have been globalized mutual efforts to overcome the serious environmental issues and increase the public perception of environmental pollution. At the United Nation Framework Convention on Climate Change(UNFCCC), the Kyoto protocol was adopted and legally obligates the developed countries to reduce their greenhouse gas emissions by at least 5.2 percent from the 1990 level by 2012. In addition to these efforts, many different international organizations are actively under discussion for environmental related issues. Currently, more than 1,000 environmental regulations are or will be set up in more than 50 countries in the world and interests of international society in environment continuously increase [Lee (2008), Chopra (2007)].

Table 1. GHG emissions and energy consumptions guideline[Korea (2010)]

Year	Type	GHG emissions allowed (unit: 1,000 tCO_2)	Energy consumptions allowed (unit: TJ)
2010	Company 1	25	500
	Workplace(Facility) 25		100
2012	Company 8	7.5	350
	Workplace(Facility) 20		90
2014	Company 5	0	200
	Workplace(Facility) 15		80

The Korean government is fully committed to low carbon green growth policies and reorganizes the environmental regulations so that it can come up to the international standard and promote the growth of environment-related industries. Table 1 describes greenhouse gas(GHG) emissions and energy consumptions guideline by Framework Act on Low Carbon, and Green Growth, which is applied to company and workplace where manager must follow two criteria. Firstly, managers of company or workplace are required to emit greenhouse gas to the maximum value which is set by guideline. Secondly, they are required to consume energy to the maximum value. Because energy consumptions are tightly linked to GHG emissions, this guideline regulates both.

According to the definition, greenhouse gases are gases in an atmosphere that absorb and emit radiation within the thermal i nfrared ra nge[Wikipedia]. T he c ontribution t o t he greenhouse effect by a gas is affec ted by both t he characteristics of the gas and its abundance. Carbon dioxide(CO_2), methane(CH_4), nitrous oxide(N_2O) and three groups of fluorinated gases, w hich are sulfur hex afluoride(SF_6), hydroflurocarbons(HFCs), and perfluorocarbons (PFCs), are the major greenhouse gases and the subject of the Kyoto Protocol, which came into force in 2005. Especially, carbon dioxide(CO_2) is the most critical one of all greenhouse gases in aspects of its abundance.

As shown in Table 2, carbon dioxide(CO_2) emissions represent 89 percent of total Korea greenhouse gas emissions, 620 million tons of CO_2 in 2007.

Table 2. GHG emissions in Korea

(Unit: million tCO_2)

	2000 2	001 2	002 2	003	2004 2	005 2	006 2	007
Gross Emissions	531	550.4 57	1.1 58	2.6	590.4 59	4.4 59	9.5	620
Rate of Increase (%)	6.4	3.7 3.	7	2	1.4 0.	7 0.	9 0.	9
Energy	438.5 45	2.9	473	481.3	489	498.5 50	5.4 52	5.4
Industrial Processes and Product Use	58.3	63.6 64	.5 68	.2	68.5 64	.8 63	.7 60	.9
Agriculture	17	16.3 16	.2	16	16.4 16	.4 15	.1 18	.4
Waste	17.2	17.6 17	.4 17	.1	16.5 14	.9 15	.4 15	.3
Forestry and Other Land Use	-37.2 -	34.6 -	33.4 -	33.7	-31.5	-32	-31.2 -	36.3
Net Emissions	493.8 51	5.8	53.8	548.8	559	562.4 56	8.4 58	3.7

Source: Ministry of Knowledge Economy

1.2 Greenhouse gas emissions in Korea

Greenhouse gas emissions increased from 531 million tons of CO_2 in 2000 to 620 million tons of CO_2 in 2007 as shown in Table 2. Average rate of increase between 2000 and 2007 is about 2.5 percent. From the base year 1990, the year when developed countries began to implement various policies to reduce their greenhouse gas emissions, to year 2007, the a verage ra te of i ncrease i s 4.3 percent. D uring t he same peri od, t he y early ene rgy c onsumptions focusing on electricity g eneration, i ndustrial pr ocesses, and tr ansportation sect or increased by 5. 6 percent on the a verage, wh ich accounts for significant effect on greenhouse gas emissions increases.

In 2007, energy sector represents 84.7 percent of total greenhouse gases emissions and becomes the largest emitter of greenhouse gases in Korea among the sectors, which also includes industrial processes and product use, agriculture, and waste. Industrial processed and product use sector represents 9.8 percent of total emissions and both of energy and industrial sect ors s how 9 4.6 pe rcent of t otal em issions. Agriculture a nd waste sectors re present 3 .0 pe rcent a nd 2 .5 percent eac h. Meanwhile the Korea n gr eenhouse gases emissions have bee n i ncreased by a bout 6 pe rcent an d t he specific sectors, which include transportation, residential, and industrial sectors, show outstanding figures in increasing rate [Kim (2009)].

Transportation sector represe nts ab out 20 percent of y early t otal green house gas em issions for e nergy sect or i n Korea. In spi te of t he i mportance, t he m ethodology t o a nalyze gree nhouse gas emissions uses a T ier 1 m ethod, the simplest one among t he m ethodologies r ecommended by IPCC guideline. Ti er 1 method ag gregates data i n a collectively manner and calculates greenhouse emissions based on total energy consumptions. Therefore, this method has a l imitation o n cal culating greenhouse gas em issions con sidering transportation infrastructure and operational environments.

This paper aims to develop a methodology of cal culating the greenhouse gas em issions for transportation sector and apply the proposed methodology to transportation sector in Korea. This paper is organized as follows. In Section 2, we rev iew the literature o n g reenhouse emissio ns m anagement system an d greenhouse em issions m anagement fo r transportation sector. Section 3 presents the methodology for calculating greenhouse gas emissions for transportation sector in Korea. Finally, concluding remarks are presented.

2. LITERATURE REVIEW

2.1 GHG emissions management system

Recently, as GHG em ission draws m uch attention worldwide, syst ems fo r m easuring a nd controlling t he GHG emissions have been politically or methodologically developed. In case of Korea, Korea Energy Economics Institute

(1996) operated and managed a total energy supply-demand statistics system, structurally compared their systems with the IEA(International Energy Agency) system, and finally suggested improvement of the total energy system. Korea Energy Economics Institute (KEEI, 2005) included an object for the national GHG emission measurement system and some considerations on operating system, which was considered as an example of the GHG emission statistics structure in Finland. KEEI report mainly suggests current states and problems of the GHG emission measurement system in Korea. Keimyung University (2006) in Korea performed a research on current conditions and issues of GHG emission of various sectors, suggested an improvement of national GHG emission measurement database. Kim, *et al.* (2006) studied management of the GHG emission in cement manufacturing industry using scenario analysis.

Table 3. GHG emissions management system in various countries

Classification	Author	Contents
Domestic	KEEI (1996)	Compared a total energy supply-demand statistics system with the IEA system, suggested improvement
	KEEI (2005)	Suggested current issues and problems of GHG emission measurement system in Korea
	Keimyung University (2006)	Suggested improvement of national GHG emission measurement database
	Kim *et al.* (2006)	Using scenario analysis, they studied the GHG emission in a cement manufacturing industry
Overseas	UNFCCC (2006)	An indicator used to announce the GHG emission and absorption of each country
	EPA (2002)	With MOBILE and MOVES, U.S. EPA office employed the GHG emission management
	Japan GHG Inventory Office (2008)	Examined the GHG emission measurement methods, and measured for the period of 1990 through 2005 in Japan
	UK Department (2007)	Introduced the national inventory report analyzing measuring methodologies of the yearly GHG emission
	Environment Canada (2007)	Published the 'national inventory report GHG sources and sinks in Canada' with the method of the GHG emission measurement
	FEA in Germany	Applied the TREMOD model for the GHG emission evaluation

With worldwide attention increasing, governments of each country or researchers interested in environment issues implemented various measurement methodologies and management systems of the GHG emission. United Nations Framework Convention on Climate Change (2006) is an indicator which is used to announce the GHG emission and absorption of each country and UN makes duty of notification. According to U.S. EPA Office of Transportation and Air Agency (2002), OAR(Office of Air and Radiation) implements the GHG emission management and OTAQ(Office of Transportation and Air Quality) measures a coefficient of emission in transportation sector using the MOBILE and MOVES models. Japan Greenhouse Gas Inventory Office of Japan (2008) examined the GHG emission measurement methods, the Center for Global Environmental Research in National Institute for Environment Studies checked emission and elimination of the GHG, indirect GHG and SO_2 from 1990 through 2005 in Japan. UK Department for Environment (2007) published the national inventory report which was written by the AEA energy & environment, analyzing measuring methodologies of the yearly GHG emission. Environment Canada (2007) published 'National Inventory Report-Greenhouse Gas Sources and Sinks in Canada' which is yearly amount of the GHG emission in Canada, is concerned with GHG emission such as CO_2, CH_4, N_2O, SF_6, PPCs and HFCs. In Germany, The German Greenhouse Gas Inventory of 1990-2004 employed a model of measuring coefficient emission by FEE(Federal Environment Agency in Germany), using TREMOD model which evaluates the amount of emission of CO_2, N_2O, CH_4, HFCs, PFCs and SF_6.

2.2 GHG emissions management for transportation sector

Though greenhouse gases are generated and managed in many sectors, in this paper we focus on the GHG emission for transportation sector. In transportation sector, researches of the GHG emission have not been made actively in Korea. Korea Energy Economics Institute (1996) studied measurement of GHG of Korea transportation, and they addressed a problem, which is insufficient statistics of gasoline and diesel oil. The statistics don't include the amount of gasoline oil consumption in military transportation and don't include and classify the diesel oil in transportation sector. For solution of this problem, they suggested a method which is accurate classification rule of GHG emission in transportation.

Table 4. GHG emissions management system for transportation sector

Classification	Author	Contents
Domestic	KEEI (1996)	Pointed out the problem of the GHG emission measurement in domestic and suggested the improvement plan
Overseas	IPCC (2006)	Suggested the standard form of measuring the GHG emission with 3 tiers of 4 transportation sectors
	UNFCCC (2006)	Measuring and managing system for the BRT of Bogota in Colombia
	EU Emission Trading	Used the method that consumption of fuel is multiplied by coefficient of emissions
	U.S. EPA (2002)	Measurement of emissions and the energy statistics for the private automobile, using RTECS
	Japan GHG Inventory Office (2008)	Through survey of investigators and mail for transportation sector, measured GHG emissions
	Environment Canada (2007)	Published the report including detailed data of transportations
	MED in Italy	MED applied the method of measuring coefficient of GHG emission, which is guidelines in Mediterranean country, to emission of transportation in Italy.

Intergovernmental Panel on Climate Change (2006) deals with many sectors, for example road, railroad, flight and ship sectors, and analyzing each sector to subdivide three tiers for GHG emissions. Tier 1 is a basic method of the GHG emission, Tier 2 and 3 are more detail methods. UNFCCC/CCNUCC (2006) has measured the emission of Colombia Bogota BRT(Bus Rapid Transit). CDM (Clean Development Mechanism), running UNFCCC/CCNUCC business, is decided that basic GHG emission is object of reduced GHG emission, and was measured the GHG emission of Bogota BRT. E. U. established GHG emissions trading for flight sector. This trading rule takes a method of measuring emissions in the way that consumption of fuel is multiplied by coefficient of emissions. The transportation statistic of U.S. EPA Office of Transportation and Air Quality (2002) is consumption data of the private automobile energy which is made by EIA(Energy Information Agency) of DOE(Department of Energy) and DOT(Department of Transportation). The consumption data of private automobile energy is made by RTECS(Residential Transportation Energy Consumption Survey). Japan Greenhouse Gas Inventory Office of Japan (2008) operates dispersed statistics management system, carrying out survey of investigators and mail for each transportation sector. Using the survey data, they measured amount of total GHG emissions for transportation. Environment Canada (2007) publishes RESD(Report on Energy supply-Demand in Canada). A detailed data for transport is made up for RESD by OEE(Office of Energy) in NRC(Natural Resources Canada) which analyzes each transport. MED(Ministry of Economic Development) in Italy applied the guideline to emission of transportation in own country. The guideline is used in Mediterranean country to measure amount and coefficient of GHG emission. Then they published the statistics year-book, contains Italian GHG emission.

In Korea, the domestic management system for GHG emissions is the lowest level, Tier 1 statistics. It is very comprehensive measuring methodology. It cannot consider a cause and direct effect of GHG emission pollution such as the characteristic of Korean driver and a traffic condition. Another problem of Korea management system is that data, amount of GHG emission of many sectors, isn't measured by correct method. These are weak point of Korean GHG management system, especially in transportation sector.

3. METHODOLOGY FOR CALCULATING GHG EMISSIONS FOR TRANSPORTATION SECTOR

Transportation sector includes road, rail, air and marine transportation. Each transportation mode is different from others in the way to calculate greenhouse gas emissions. In general, calculation methods of greenhouse gas emissions are based on Intergovernmental Panel on Climate Change(IPCC) guidelines which are used for reporting nation's greenhouse gas emissions to United Nations Framework Convention on Climate Change(UNFCCC), an international treaty to begin to consider what can be done to reduce global warming and to cope with whatever temperature increases are inevitable.

This section reviews the general methodologies of calculating the greenhouse gases emissions for transportation sector and describes the methodology for greenhouse gases emissions analysis in collectable and measurable ways. The proposed methodology considers nation's transportation infrastructure and operational environments. For example, transport modes and vehicle types used in Korea are considered.

3.1 General methodology of calculating the greenhouse gases emissions for the transportation sector

The most common simple methodological approach to calculate GHG emissions is to combine information on the extent to which a human activity takes place (called activity data or A_i from Equation (1)) with coefficients which quantify the emissions or removals per unit activity(called emission factors or EF_i from Equation (1))[IPCC (2006)]. The basic equation is as follows:

$$EM = \sum_{i=1}^{n} A_i \times EF_i \tag{1}$$

where

EM : CO_2 emissions (kg)

A_i : Activity Data (Fuel consumption) (TJ)

EF_i : Emission Factor (kg/TJ)

i : Type of Fuel

Table 5. CO_2 emissions by major fuels

Fuel Unit		IPCC Carbon Emission Factors(unit: t/TOE)	Energy Conversion Factors(unit: TOE)	CO_2 emissions per unit (t/L)
Motor Gasoline	L 0	.783	0.740	0.783*0.740*44/12
Diesel Oil	L	0.837	0.845	0.837*0.845*44/12
LPG(Liquefied Petroleum Gases)	L 0	.713	1.380	0.713*1.380*44/12

For example, in the transportation sector fuel consumption would constitute activity data, and mass of carbon dioxide emitted per unit of fuel consumed would be an emission factor. From Equation (1), emission factor(EF_i) is calculated by product of carbon emission factors which are set by IPCC guidelines and energy conversion factors which are provided by Energy Act and used for converting Ton of Oil Equivalent(TOE) to unit of measure by fuel, and the ratio of mass between carbon atoms to CO_2 which is used for converting carbon atoms to CO_2 when combustion occurs. The ratio of mass between carbon atoms to CO_2 is set to 44/12 determined by chemical formula. For example, CO_2 emissions by major fuels used in cars is shown as Table 5. CO_2 emissions per unit for vehicles using motor gasoline are product of 0.783 and 0.740, and 44/12.

3.2 Road Transportation

Road transportation represents the single largest source of GHG emissions in the transportation sector, accounting for more than 50 percent of total transportation emission in most countries. CO_2 emissions for road transportation can be calculated based on fuel consumptions used by Light Duty Vehicle(LDV)(Ex. passenger car, light truck *et al.*), Heavy Duty Vehicle(HDV)(Ex. Bus, Heavy Truck *et al.*), and motorcycle. Basic CO_2 emission calculating method is the same as Equation (1), product of activity data like fuel consumptions and emission factor. In other words, yearly CO_2 emissions for road transportation can be analyzed by combination of total fuel consumptions of all vehicles during one year and emission factors by fuel . Activity data to analyze CO_2 emissions for road transportation is simply fuel consumptions and it requires more detailed activity data for accurate CO_2 emissions analysis. To analyze the activities in detail, IDEF0 model is used and the key components of model are fuel injected, passenger and freight loaded, weight of vehicle, engine and fuel types of vehicle, CO_2 emissions generated by driving a vehicle, distance travelled, driving style of driver, and road surface as shown as Figure 1. For example, fuel types of vehicle are motor gasoline, diesel oil, and Liquefied Petroleum Gases(LPG) and engine type means the technology level of vehicle to reduce CO_2 emissions by using emerging and advanced technologies.

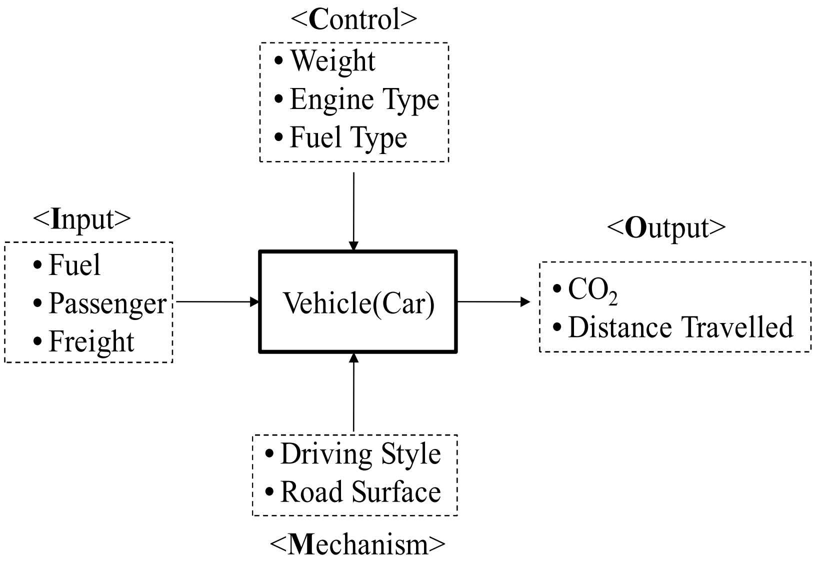


Figure 1. Analysis of Road Transportation using IDEF0

Total travel distance of vehicle can be calculated by product of total vehicles registered and daily distance travelled, and days travelled as shown as Equation (2). We assumed that the value of day travelled equals 365. And total fuel consumed of vehicles can be analyzed by using the average fuel efficiency as shown as Equation (3).

$$D_{ij} = TV_{ij} \times VM_{ij} \times 365 \quad (2)$$

where

D_{ij} : Yearly Total Distance Travelled of Vehicle j using Fuel Type i (km)

TV_{ij} : Total Vehicles Registered of Vehicle j using Fuel Type i

VM_{ij} : Average Daily Distance Travelled of Vehicle j using Fuel Type i (km)

$$F_{ij} = D_{ij} / FE_{ij} \quad (3)$$

where

F_{ij} : Yearly Total Fuel Consumed of Vehicle j using Fuel Type i (L)

FE_{ij} : Average Fuel Efficiency of Vehicle j using Fuel Type i (km/L)

In the way activity data are expanded, emission factor can be segmented into Equation (4) as explained in Table 5. Therefore CO_2 emissions calculation formula for road transportation can be transformed into Equation (5) using Equation (3) and Equation (4).

$$EF_{ij} = CEF_{ij} \times ECF_{ij} \times \frac{44}{12} \quad (4)$$

where

CEF_{ij} : IPCC Carbon Emission Factors of Vehicle j using Fuel Type i (t/TOE)

ECF_{ij} : Energy Conversion Factors of Vehicle j using Fuel Type i (TOE)

$$EM = \sum_{i=1}^{n} \sum_{j=1}^{m} \left\{ \left[\frac{TV_{ij} \times VM_{ij} \times 365}{FE_{ij}} \right] \times \left[CEF_{ij} \times ECF_{ij} \times \frac{44}{12} \right] \right\} \quad (5)$$

3.3 Rail Transportation

Railroad locomotives have three types: diesel, electric, and steam. Diesel locomotive uses diesel engine and electric locomotive uses electricity through the pantograph, a device that collects electric current from overhead lines for electric trains or trams. Steam locomotive doesn't be working for general logistic use but for limited use like sightseeing.

CO_2 emissions from railroad locomotives can be calculated by using total fuel consumptions in general. Basic method of CO_2 emissions analysis, which is based on Tier 1 method by IPCC guideline, is the same as Equation (1). But advanced method has to consider more detailed variables including total fuel consumptions. For example, type of locomotives, distance travelled by locomotive, and time between origin(city) and destination(city) i.e. operating hours can be extra variables for CO_2 emissions calculation. Two types of trains has been working, KTX using electricity as fuel and Saemaeul or Mugunghwa using diesel oil as fuel, in Korea. Most people would take KTX because of speed and availability of train.

CO_2 emissions for rail transportation can be analyzed by modifying the Equation (5). Activity data from Equation (5) can be modified into Equation (6) and emission factor is the same as Equation (5). Therefore CO_2 emissions calculation formula for rail transportation can be transformed into Equation (7) using Equation (5) and Equation (6).

$$A_i^R = \sum_{j=1}^{m} F_{ij}^R = \sum_{j=1}^{m} \left[\frac{D_{ij}^R}{FE_{ij}^R} \right] \tag{6}$$

where

A_{ij}^R : Activity Data by Fuel Type i (L)

F_{ij}^R : Yearly Total Fuel Consumed of Locomotive j using Fuel Type i (L)

D_{ij}^R : Yearly Total Distance Travelled of Locomotive j using Fuel Type i (km)

FE_{ij}^R : Average Fuel Efficiency of Locomotive j using Fuel Type i (km/L)

$$EM^R = \sum_{i=1}^{n}\sum_{j=1}^{m} \left\{ \frac{D_{ij}^R}{FE_{ij}^R} \times \left[CEF_{ij} \times ECF_{ij} \times \frac{44}{12} \right] \right\} \tag{7}$$

3.4 Air Transportation

CO_2 emissions from air transportation in Korea are the third or fourth source of emissions in the transportation sector. CO_2 emissions from air transportation will increase because of growing aircraft travel demand. CO_2 emissions from aircraft can be calculated by using total fuel consumptions like road or rail transportation. Basic method of CO_2 emissions analysis is thus the same as Equation (1). But advanced method has to consider more detailed variables including total fuel consumptions. For example, type of aircraft, landing/take-off cycles(LTOs) of aircraft, and available seat miles(ASM) which is a measure of an airline flight's passenger carrying capacity can be extra variables for CO_2 emissions calculation. Korea Airlines and Asiana Airlines are more popular than other foreign airlines in Korea. Korea Airlines has 127 aircrafts and Asiana Airlines has 67 aircrafts in 2010.

CO_2 emissions for air transportation can also be analyzed by modifying the Equation (5) like rail transportation. However numbers of landing/take-off must be considered to reflect the characteristics of air transportation. Therefore CO_2 emissions calculation formula for air transportation can be transformed into Equation (8).

$$EM^A = \sum_{i=1}^{n}\sum_{j=1}^{m} \left\{ \left[\frac{D_{ij}^A}{FE_{ij}^A} \times \right] \times \left[CEF_{ij} \times ECF_{ij} \times \frac{44}{12} \right] + LT_{ij}^A \times \left[CEF_{ij}^A \times ECF_{ij} \times \frac{44}{12} \right] \right\} \tag{8}$$

where

D_{ij}^A : Yearly Total Distance Travelled of Aircraft j using Fuel Type i (miles)

FE_{ij}^A : Average Fuel Efficiency of Aircraft j using Fuel Type i (miles/L)

LT_{ij}^A : Yearly Landing/Take-off of Aircraft j using Fuel Type i

CEF_{ij}^A : Carbon Emission Factors when Landing/Take-off of Aircraft j using Fuel Type i (t/TOE)

3.5 Marine(Ship) Transportation

CO_2 emissions from marine transportation in Korea are the second source of emissions in the transportation sector. Method of CO_2 emissions analysis in marine transportation has to consider type of ship and type of engine including type of fuel. CO_2 emissions calculation formula for marine transportation can be transformed into Equation (9) using Equation (7).

$$EM^M = \sum_{i=1}^{n}\sum_{j=1}^{m}\left\{\frac{D_{ij}^M}{FE_{ij}^M}\times\left[CEF_{ij}\times ECF_{ij}\times\frac{44}{12}\right]\right\} \tag{9}$$

where

D_{ij}^M : Yearly Total Distance Travelled of Ship j using Fuel Type i (miles)

FE_{ij}^M : Average Fuel Efficiency of Ship j using Fuel Type i (miles/L)

4. CONCLUSIONS

This paper aims to develop a methodology of calculating the greenhouse gas emissions for transportation sector and apply the proposed methodology to transportation sector in Korea. We represented the methodology of calculating the greenhouse gas emissions, especially CO_2, for transportation sector and apply this methodology to transportation sector in Korea. Transportation sector is one of the most critical sectors in view of environmental pollutions and represents about 20 percent of yearly total CO_2 emissions for energy sector in Korea. In spite of the importance, the methodology to analyze CO_2 emissions uses a Tier 1 method, the simplest one among the methodologies recommended by IPCC guideline. For this reason, we developed the methodology of calculating the CO_2 emissions more accurately for the transportation sector.

CO_2 emissions analytic methodology can analyze the amount of CO_2 emissions for each transport mode and integrate it through transportation environments. It can also provide the decision-making tool for government to make the transportation polices. If the statistics data can be provided sufficiently and CO_2 emissions for each transport mode can be calculated more accurately, Korean government will effectively manage the transportation sector and improve national eco-environment.

ACKNOWLEDGEMENTS

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No.2009-0085893).

6. REFERENCES

Chopra, S. and Meindl, P. (2007). *Supply Chain Management: Strategy, Planning and Operation*, Pearson Education.

Entec UK (2008). Draft Guidelines on Monitoring, Reporting and Verification(MRV) of Greenhouse Gas Emissions for the Aviation Sector under the European Union Emissions Trading Scheme(EU-ETS).

Entec UK (2008). European Commission MRV Guidance for aviation in the EU ETS-Annual Emissions Draft Technical Report.

Environment Canada (2007). National Inventory Report-Greenhouse Gas Sources and Sinks in Canada 1990-2005, Green house Gas Division.

Federal Environmental Agency (2006). National Inventory Report for the German Greenhouse Gas Inventory 1990-2004.

German Federal Environmental Agency: Federal Environmental Agency Homepage. http://www.fea.gov.ae/

Intergovernmental Panel on Climate Change (2006). 2006 IPCC Guidelines for National Greenhouse Gas Inventories.

Japan Greenhouse Gas Inventory Office (2008). National Greenhouse Gas Inventory Report of Japan, Ministry of Environment.

Keimyung University(2006). Building National GHG Emission Statistics Database for Establishing National Strategic Road-map and Research of Optimum Operation Method.

Kim, H. S. and Kang, H. J. (2006). Analysis of the Greenhouse Gas Reduction Scenarios in the Cement Manufacturing Industry, *Journal of Korean Society for Atmospheric Environment, 6,* 912-921.

Kim, S. R. (2009). Issue Report: Introducing Directions on CO_2 Taxes for Green Growth.

Korea (2010). Framework Act on Low Carbon, Green Growth.
Korea Energy Economics Institute (1996). Energy Supply-Demand Statistic Syste m and Amount of Change Standard Research.
Korea Energy Economics Institute (2005). GHG Inventory Measurement and Statistic System Improvement.
Korea Energy Economics Institute (2009). Yearbook of Energy Statistics 2009.
Korea Energy Economics Institute (2009). Aviation Statistics 2009.
Korea Transportation Safety Authority (2008). Survey on the Mileage of a Car.
Lee, J. (2008). The Advent of the Era of Green Growth, Samsung Economic Research Institute CEO Information.
Lee, K. (2008). Era of Climate Change and Environmental Management: Changes in Global Markets - Challenges and Opportunities, IBM Big Green 2.0.
Ministry of Knowledge Economy (2009). Energy Consumption Survey 2008.
Ministry of Knowledge Economy. A Study on Framework for Greenhouse Gas Inventory Development.
Ministry of Land, Transport and Maritime Affairs (2009). Statistical Yearbook of Land, Transport and Maritime.
Ministry of Land, Transport and Maritime Affairs (2010). Statistical Yearbook of Traffic Volume.
UK Department for Environment (2007). UK Greenhouse Gas Inventory, 1990 to 2005-Annual Report for Submission under the Framework, Convention on Climate Change, Food and Rural Affairs.
UNFCCC/CCNUCC (2006). Metho dology fo r Bu s Rap id Tran sit Projects(BRT Bogo ta, C olombia : Tran sMilenio Phase II to IV).
U.S. Environmental Protection Agency (2007). Inventory of U.S. Greenhouse Gas Emission and Sinks : 1990-2005.
U.S. EPA Office of Transportation and Air Quality (2002). EPA's Plan for MOVES : A Comprehensive Mobile Source Emission Model, Assessment and Standards Division.
UK Greenhouse Gas National Inventory: UK GHG Nation Inventory Home Page. http://www.ghgi.org.uk/
Wikipedia Home Page. http://en.wikipedia.org/wiki/Greenhouse_gas/

BATCHING MANAGEMENT IN WAREHOUSES FOR COMPANIES WITH HIGH RETURN FLOWS - DESIGN AND CATEGORIZATION OF MODELS AND APPROACHES

Susanne Wruck[1,2], Iris F.A.[1] Vis, and Jaap Boter[2,3]

[1]Department of Information Systems and Logistics, [2]Department of Marketing
Vrije Universiteit Amsterdam
De Boelelaan 1105
NL-1081 HV Amsterdam, The Netherlands
e-mail: swruck@feweb.vu.nl, ivis@feweb.vu.nl, jboter@feweb.vu.nl

[3]Department of Book Studies
University of Amsterdam
Oude Turfmarkt 141
NL-1012 GC Amsterdam, The Netherlands
Email: j.boter@uval.nl

Abstract: With the rise of Internet sales, retailers have witnessed increased return flows, as website pictures and specifications alone are often insufficient for customers when deciding on the right size, color or suitability of the product. Since an efficient inclusion of the return flow can lead to an improved inventory management and a higher utilization of warehouse capacities, it is most important to develop appropriate optimization models. The problem of composing a set of products into efficient batches has been addressed in literature for forward flows before, but the increased backward product flows require a new consideration. To address this problem, we design and classify optimization models into deterministic, multistage stochastic and online approaches. We discuss the implications of each model and analyze their scope of application for integrated forward and return flows by proposing several optimization objectives. We conclude with an overview of possible solution methods for each model.

1. INTRODUCTION

Order picking is known to be the most labor intensive process in warehouse operations (de Koster et al., 2007). Considerable cost savings can be achieved, if the included procedures, such as routing and batching policies are designed most efficiently. Due to a continuous growth of Internet sales (Bellman et al., 1999), companies are increasingly faced with the challenges of integrated forward and return flows. Pictures and descriptions at websites are sometimes insufficient for customers when making their purchase decision. Return rates up to 35% are therefore common in many companies. Efficient processing and incorporation of the return flow will lead to an improved inventory management and a higher utilization of warehouse capacities.

Batching is the method of grouping a set of orders into a number of subsets, each of which can then be retrieved by a single tour (de Koster et al., 2007). It is often better to group several small orders into a larger amount of work, to realize an increase in labor efficiency (Gibson and Sharp, 1992). Gademann and van de Velde (2005) have analyzed the complexity of the batching problem and highlighted the need for effectual algorithms to solve it. Since then, the batching problem for forward distribution flows has been addressed in research studies (Ho et al., 2008, Henn et al., 2009, Hwang and Him, 2005, and Won and Olafsson, 2005). Now, the increased reverse product flows require a new consideration of the problem, since more efficient batching policies can be derived when returns are directly incorporated. For the restoring process of products there is often a more flexible time restriction required, so that restoring activities can be incorporated into batches, in which the capacity (labor and transport) is not completely utilized by picking activities. If in these situations additionally returns are included in batches while the picking process is not affected, this can offer new opportunities in the design of beneficial policies.

In previous research highly efficient policies were identified to construct batches for the forward product flow. Those policies could be implemented to separately process the return flow, but we suggest a batching management with integrated forward and return flow handling, since it can lead to higher warehouse operation performance (Wruck et al., 2010). Since order picking is an essential link in the supply chain and is the major component of warehousing (Petersen, 1997), a focus on the optimization of integrated forward and return handling in the batching management is of high interest. We design three optimization models for the batching problem with integrated forward and return flows. First, in section 3.1 we consider a deterministic offline approach, which will turn out to be a suitable model to derive tactical solutions, i.e., a general batching policy matching the warehouse situation in practice. That means that for a specific warehouse, stock size, average number of orders, and return rate, a certain batching procedure can be recommended (average batch size, length of tour start intervals, and the number of included returns), that leads to high performance in terms of processing times and overall travel distance. Second, we present in section 3.2 another offline approach, a multistage stochastic problem formulation. This artful, but complex model can be used when predictions on future events are supposed to be incorporated. Third, an online approach will be proposed in section 3.3, in which the problem is considered at a certain time step and a solution, i.e., the shape of the batch is determined only based on information about the past. For each model we will introduce suitable performance measures and propose several optimization objectives. In section 4 we illustrate the implications of each model and analyze their scope of application. Moreover, an overview of promising solution techniques for each model is presented.

2. PROBLEM DESCRIPTION

We consider the problem of efficiently dividing a set of jobs (i.e. picking and restoring activities) that can be conducted each by one employee in a warehouse. We propose a particular focus on companies with high returning flows, that is customers do not only order products but returns also arrive in the warehouse and are supposed to be reintegrated in the stock as soon as possible, so that the inventory can be managed most effectively (for instance in terms of replenishment). Decisions that have to be made concerning the creation of a batch, i.e., the batching strategy, are (1) the overall number of jobs that are contained; (2) the amount of restoring activities that is contained; and (3) the product selection based on their proximity in the storage.

The batching strategy is closely related to and its performance dependent on the underlying routing policy (Albareda-Sambola et al., 2009). Hence, a fixed routing and storage location policy will be used in our models, so that the length of a route is known for each set of jobs. Which routing and storing method is used therefore does not affect the solution of the batching problem. Moreover, our models can be implemented in warehouses with one depot, meaning that returns and customer orders arrive at the same point and picked product have to be delivered to that point.

We consider the following setting. Customers order products to be picked by an employee from a specific storage location. At the moment the order arrives at the warehouse the lead time of the product is dependent on the processing within the warehouse. Moreover, customers return products that have to be restored. Customers returning products are not interested in the time to restore the product, so the time restriction for a restoring activity is more flexible than the time restriction for waiting customers. However, an effective restoring system can strengthen the inventory and service management of a firm. For instance, in a library it is desirable that books are quickly available again for a new customer after their return. The warehouse of a catalog company can organize its replenishment more efficiently, if returned products are integrated in the inventory system as soon as possible. Hence, we restrict the time window for the restoring process, too. All returned products have to be restored at the latest at the end of day.

We denote the set of all requested and returned products as the set of jobs to be conducted. Employees who start a tour in order to pick and/or return a set of products follow a particular routing schedule and have a constant pace. A job that occurs after an employee's tour has started cannot be included in the tour anymore and has to be conducted in a later tour or by another employee. The number of employees who can conduct jobs and the number of jobs that can be conducted in one tour, i.e., the transport capacity, are limited. We propose models to approach the batching problem on different decision levels in order to determine the best batching strategy for specific warehouse situations, whereas the latter is determined by (1) the overall number of jobs, (2) the rate of contained returns, and (3) the number of employees available. Our aim is to demonstrate and explain different views on the batching problem and to indicate solution approaches at different decision levels. In doing so, we want to enrich the discussion of batching management for integrated forward and return product flows by formulating suitable optimization models.

3. MODEL DEVELOPMENT

The highest decision level at which batching is of interest is the tactical level. Tactical decisions typically concern the dimensions of resources, i.e., the number of employees, the determination of a layout and a number of organizational issues (Rouwenhorst et al., 2000). Our first model, the deterministic optimization model, we propose for practical problems, in which batching procedures, i.e., the transport capacity and the number of employees, are supposed to be determined in order to achieve the desired warehouse operation performance. As long as the operation within the warehouse and the job flow do not change, these general batching procedures are tactical decisions that only have to be reconsidered medium-term. The second approach is the formulation as a multistage stochastic optimization program. Similar to the deterministic model this problem formulation will cover the entire time horizon and can therefore be set at the tactical decision level. This model incorporates information about future events with the help of their probabilistic characterization; thus independent from historical data. The third approach is designed to be implemented online, meaning we react with a solution (tour start yes or no, which products are contained) on a real current situation at a certain time step. This approach can be used to refine general solutions identified with approaches at the tactical level. The difference in modeling with integrated forward and flows is that those two types of jobs have to be included in the model separately, although a simultaneous optimization is targeted. The constraints to limit the batch size and to formulate the tour length can be modeled for mixed batches, but other restrictions, for instance the allowed time window for processing a job require a distinction between picking and restoring activities. The performance of the delivery process (i.e. forward) should not be affected by included returns. So restrictions are required to facilitate a delivery of ordered products on time, while returns might be postponed until a more suitable batch can be created. For this purpose we will differentiate between orders and return jobs in our models and we formulate restrictions to deal with each type appropriately.

3.1 Deterministic Optimization Model

This approach is dedicated to the determination of a general batching strategy for specific warehouse settings. That is, the average batch size, average length of tour start intervals, and which returns should be included are supposed to be decided for a practical situation, i.e., a storage allocation system, a certain stock size, and forward and return flows. The model can also be implemented when related tactical decisions are supposed to be made, for instance the number of employees and transport capacity needed, in order to achieve a desired batching policy performance.

Approaching the batching problem in this way implies to find optimal solutions in several calculations, i.e. a partitioning of all jobs into batches, for specific settings with the aim to identify batching strategies for classes of settings. That means, several similar settings (similar stock size, similar layouts, and similar return rates) build a setting class, for which a suggestion for a batching strategy can be derived by identifying patterns in the corresponding solutions that all of these settings share.

To develop the deterministic model let $\mathcal{T} = \{0, \dots, T\}$ be a discrete time horizon. Jobs (order or return) can arrive at any time step in $\mathcal{T}$ and a tour of an employee can start at each time step. Moreover, we define the set of all products by $\mathcal{N} = \{1, \dots, N\}$. Two kinds of jobs that may occur are considered. Products are either ordered or returned by customers. Hence, we define the set $\mathcal{S} = \{0,1\}$, whereat 0 denotes that a product is ordered and 1 denotes that a product is returned.

With the help of these notations we interpret each occurring job *j* as an element of the set

$$\mathcal{J} \subset \mathcal{N} \times \mathcal{T} \times \mathcal{S}. \tag{1}$$

In the following we interpret a job $j \in \mathcal{J}$ as the 3-tuple $j = (n, t, s)$, whereby $n \in \mathcal{N}$, $t \in \mathcal{T}$, and $s \in \mathcal{S}$. We introduce two types of variables that will be used to describe the procedure of a tour composition. Let $\mathcal{M} = \{1, \dots, M\}$ be the number of employees that can conduct tours. First, for every time step $t \in \mathcal{T}$ and $m \in \mathcal{M}$ we define the binary variable $r^{t,m}$ to formulate an employee's start of a tour by

$$r^{t,m} = \begin{cases} 0, & \text{no tour start of employee } m \text{ in time step } t \\ 1, & \text{tour strat of employee } m \text{ in time step } t. \end{cases} \tag{2}$$

Second, we describe the inclusion of a job $j \in \mathcal{J}$ in a tour starting at time step $t \in \mathcal{T}$ by using the binary variable $inc_j^{t,m}$.

$$inc_j^{t,m} = \begin{cases} 0, & \text{job } j \text{ is not included in the tour of employee } m \text{ starting at } t \\ 1, & \text{job } j \text{ is included in the tour of employee } m \text{ starting at } t \end{cases} \tag{3}$$

After introducing the notations, we will define several objectives for the deterministic optimization model. After that, we will formulate the necessary constraints.

3.1.1 Minimize the overall tour length

The length of a tour depends on the number of jobs that are contained and the storage location of each product. Let $J^{t,m} = \{j_1, \dots, j_K\} \subset \mathcal{J}$ be a set of jobs to be executed by employee m in one tour at time *t*, that is $inc_{j_k}^{t,m} = 1$ for all $k = 1, \dots, K$. The resulting length of this tour will be denoted by $L_{J^{t,m}}$.

To formulate the objective of minimizing the total travel distance for all employees within the overall time period we first describe the total number *H* of tours conducted by all employees by

$$H = \sum_{m=1}^{M} \sum_{t=1}^{T} r^{t,m} \tag{4}$$

Let $t_1, \ldots, t_H$ denote the time steps in which tours start and m_h be the employee who conducts the tour starting at t_h. Then, the objective is given by

$$min \left\{ \sum_{h=1}^{H} L_{J^{t_h m_h}} \mid \bigcup_{h=1}^{H} J^{t_h m_h} = \mathcal{J} \right\} \tag{5}$$

3.1.2 <u>Minimize the average processing time</u>

We formulate an objective in order to minimize the average lead time of an ordered product, i.e. the time-lag between the arrival of the order at the warehouse and its delivery to the depot. This target will constitutionally lead to a reduction of the number of returns in a batch. Hence, it must be taken into account that a constraint is needed in order to exact a restoring of returns at certain time steps in order to find solutions that incorporate return flow handling. An appropriate constraint will be constructed in the next section. We denote by $p \in \mathbb{R}_+$ the pace of employees moving in warehouse. Thus, the duration of a tour of employee *m* starting at time step *t* and consisting of the jobs $J^{t,m} \subset \mathcal{J}$ is given by

$$D_{J^{t,m}} = \frac{L_{J^{t,m}}}{p} \qquad \forall\, j \in \mathcal{J}\colon\ inc_j^{t,m} = 1 \tag{6}$$

The lead time of a product $j = (n, t', 0)$ that has been ordered at $t' \leq t$ and is contained in this tour can be formulated by

$$W_j = t - t' + D_{J^{t,m}} \qquad \forall\, j \in \mathcal{J}\colon\ inc_j^{t,m} = 1. \tag{7}$$

Let $|\mathcal{J}_{ord}|$ be the number of requested products within the time horizon. The objective of minimized average processing time of ordered products is then given by

$$min \left\{ \frac{1}{|\mathcal{J}_{ord}|} \sum_{j \in \mathcal{J}_{ord}} W_j \right\}. \tag{8}$$

3.1.3 <u>Minimize lateness</u>

In some settings promises concerning delivery times are made. Assume that a product has to be available at the depot until time t_2, if the order arrived before time $t_1 < t_2$. This can be expressed by the following relationship.

$$\forall\, j = (n, t', 0) \in \mathcal{J} \text{ such that } t' \leq t_1 \quad \Longrightarrow \quad t' + W_j \leq t_2 \tag{9}$$

A third objective might therefore be to minimize the lateness of delivery. We introduce another binary variable that describes whether lateness occurs. For $j = (n, t', 0) \in \mathcal{J}$ let

$$late_j = \begin{cases} 0\,, & t' \geq t \\ 0, & t' \leq t,\ t' + W_j \leq t_2 \\ 1, & t' \leq t,\ t' + W_j \geq t_2 \end{cases} \tag{10}$$

With the help of this variable the objective of a minimized number of late deliveries can be formulated by

$$min \left\{ \sum_{j \in \mathcal{J}} late_j \right\}. \tag{11}$$

3.1.4 Constraints

The necessary constraints depend on the objective function that has to be optimized. We model several restrictions and explain for which objectives they are needed. Two constraints that are needed for the definition of the inclusion variable in (3) have to be incorporated. This variable can only be 1, if employee m starts a tour at time step t (12) and a job can only be included, if it occurs before a tour starts (13). Moreover, every job has to be executed only once, which leads to the constraint (14).

$$inc_j^{t,m} \leq r^{t,m} \qquad \forall m \in \mathcal{M}, \forall\, t \in \mathcal{T} \tag{12}$$

$$inc_j^{t,m} \leq 0 \qquad \forall m \in \mathcal{M}, j = (n, t', s), s \in \mathcal{S}, t \in \mathcal{T}\,, t' > t \tag{13}$$

$$\sum_{m=1}^{M} \sum_{t=1}^{T} inc_j^{t,m} = 1 \qquad \forall j \in \mathcal{J} \tag{14}$$

Another necessary constraint is that a new tour of employee m cannot start as long the employee's previous tour has not ended. Let t_1 be the time step in which the previous tour has started. This leads to the restriction

$$r^{t,m} = 0 \;\; \forall\, t \in \mathcal{T}\!: \; t_1 < t \leq t_1 + D_{J^{t_h},m}\,, \;\; \forall\, m \,\in \mathcal{M}. \tag{15}$$

If the objective is the minimization of the total walking distance as described in (5), a constraint must be included that limits the delivery time for products to the depot. Hence, we add the maximum lead time allowed W_{max}^{ord} between the arrival of a product at the warehouse and its delivery to the depot, whereby we assume that this time period is equally long for every ordered product. The corresponding limitation can be express by

$$W_j \leq W_{max}^{ord} \qquad \forall\, j = (n, t', 0) \in \mathcal{J}. \tag{16}$$

Moreover, we incorporate a similar constraint for return jobs by requiring

$$W_j \leq W_{max}^{ret} \qquad \forall\, j = (n, t', 1) \in \mathcal{J}, \tag{17}$$

whereby W_{max}^{ret} denotes the maximum time allowed between the arrival of a returned product and its restoring in the warehouse.

For settings in which delivery promises are incorporated, either as a target or as a constraint, the above mentioned restriction is only required for orders not already covered by the lateness restriction.

An adequate constraint to control for exceeded lead times can be added. That means, we allow for lateness, but we target a solution with a minimal number of those. This limitation can be modeled as follows. Let $late_{max}$ be the maximum number of lateness's suffered. Then we can limit late deliveries in requiring

$$late_j \leq late_{max} \qquad \forall\, j = (n, t, 0) \in \mathcal{J}. \tag{18}$$

In all settings a limitation of the batch size is necessary. The number of products that can be picked or returned in one tour should not exceed a certain value b_{max}. To account for that we denote by $|J^{t,m}|$ the number of jobs $J^{t,m} \subset \mathcal{J}$ that are contained in a tour of employee m starting at time t and we formulate the restriction

$$|J^{t,m}| \;\leq b_{max} \quad \forall t \;\in \mathcal{T}, \forall m \in \mathcal{M}. \tag{19}$$

Which products we should check first in order to decide whether they are included in the batch or not depends on the model that is used. In the deterministic model a specific order is not of interest, since all jobs are known in advance and the objective is the optimal partitioning along the entire time horizon. In the next section we discuss an online approach in which at each time step a number of jobs is available. In this case the selection procedure should be organized in a way that jobs arrived earliest are checked first to be included, since those are more likely to be urgent.

3.2 Multistage Stochastic Optimization Model

Next to a deterministic formulation, the batching problem can be approached with a stochastic optimization model. The most sophisticated but also most complex option to manage batching is to make decisions for each time step based on information about realized events in the past, while accounting for expected events in the future. Figure 1 illustrates the procedure.

Future events are often characterized as scenario trees (Heitsch and Rőmisch, 2009, and Consigli and Dempster, 1998). That is, a finite number of different outcomes is defined for each time step and their probabilities are estimated. Each path within the tree can then be described with a certain probability and expectation values about the future outcome can be incorporated in the optimization. In doing so the solution might become more consistent than a deterministic solution, since it is independent from historical data. Unlikely realizations might not be covered by the deterministic solution if they do not occur in the data set, whereas in a stochastic model all possible outcomes and their probabilities are taken into account. The well-known drawbacks of stochastic modeling are the complexity and high-dimensionality

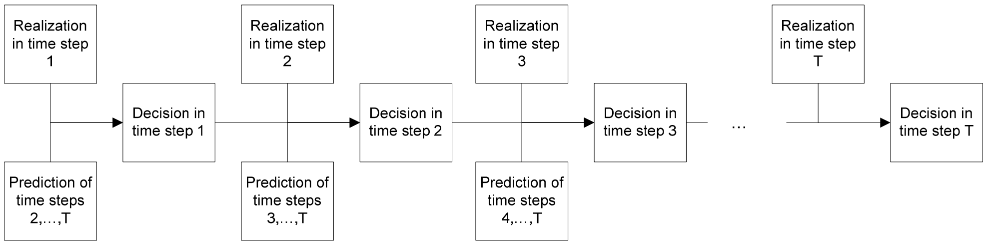


Figure 1: Problem solution with a stochastic approach

due to the complex description of the stochastic processes. The incorporation of stochastic characterizations in online approaches is therefore only a theoretical option, since computing times normally forbid an implementation in practice and approximations with reduced complexity that still reflect reality are challenging to develop.

For that reason we develop a multistage stochastic formulation that, similar to the deterministic model, simultaneously covers the entire time horizon and can therefore be implemented to make decisions concerning general policies. Multistage stochastic models involve a stochastic process and a corresponding decision process, in which each decision is dependent on past events (realizations of the stochastic process and decisions). Unlike deterministic models, the target is to optimize the expected value of the objective function, which depends on the stochastic process and the decision process. In doing so, uncertainty about future events is taken into account at each step and we obtain optimization models that reflect reality better than deterministic models.

The stochastic influence in batching problems is the expected job flow, while the elements of the decision process contain all decisions that have to be made at each time step (tour start yes or no, batch size, batch composition).

To develop the model let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, whereby Ω is a sample space, $\mathcal{F}$ a σ-algebra and $\mathbb{P}$ a probability function with $\mathbb{P}: \mathcal{F} \longrightarrow [0,1]$. Let $\{\xi^t\}_{t=1,\dots,T}$ be the flow of arriving jobs as the stochastic process defined on $(\Omega, \mathcal{F}, \mathbb{P})$. That means that the random variable $\xi^t = \{j_1^t, \dots, j_l^t\}$ describes a set of jobs that arrive at time step $t \in \mathcal{T}$. Further, we denote by $\xi^{[t]} = (\xi^1, \dots, \xi^t)$ the vector of random variables until time step *t*.

In the decision process we merge all decision variables for one time step *t* (i.e. $r^{t,m}$ and $inc_j^{t,m}$ $\forall\, m \in \mathcal{M}, j \in \mathcal{J}$) in a vector x^t. So, the process of decision variables is denoted by $\{x^t\}_{t=1,\dots,T}$. Analog to the objective of minimizing the overall travel distance in the deterministic formulation (5) we can define the objective to minimize the *expected value* of the overall travel distance by

$$min\left\{\mathbb{E}\left[\sum_{h=1}^{H} L_{J^{t_h,m_h}}\left(\xi^{[t_h]}, x^{t_h}\right)\right] \;\mid\; \bigcup_{h=1}^{H} J^{t_h,m_h} = \mathcal{J} \;\; \mathbb{P}-a.s.\right\}, \tag{20}$$

whereby $h = 1, \dots, H$, again, denotes these time steps in which tours start, and the constraint concerning the partitioning of $\mathcal{J}$ is required $\mathbb{P}$-almost sure[1]. Tour length $L_{J^{t,m}}$ now becomes a random variable, since it is dependent on the random variables of the job flow until time step t $\left(\xi^{[t]}\right)$ and the decision in time step t (x^t).

Analog to this example, other objective functions, such as average processing time as described in (8) and the number of late deliveries as proposed in (11) can be adjusted and set to be optimized in multistage stochastic models.

Similarly, constraints that have been developed for the deterministic model have to be adapted. The lead time W_j for instance, as described in (7) also becomes a random variable, since it is dependent on the random variable tour length. The constraint of bounding the processing time of an ordered job is therefore also required to hold $\mathbb{P}$-almost sure and can be formulated as follows.

$$W_j\left(\xi^{[t']}, x^{t'}\right) \le W_{max}^{ord} \qquad \forall\, j = (n,t,0) \in \mathcal{J} \quad \mathbb{P}-a.s. \tag{21}$$

[1] An event is called $\mathbb{P}$-almost sure, it its incidence has probability 1.

In this paper, referring to the example of processing time limitation (21), we exclude the formulation of the remaining constraints that are needed to complete the multistage stochastic model.

A multistage stochastic problem formulation allows a more realistic consideration of the batching problem than the deterministic model. In the deterministic model decisions at each time step are made based on information of the job flows of the overall time horizon, whereas in the stochastic model decisions are only linked to events in the past and the expected outcome is optimized.

3.3 Online Approach

Considering batching as a decision problem in a certain time step requires formulating an online approach, in order to find the best solution for the current situation. In this case, information about incoming flows is not known in advance. Rather, the problem is solved at each time step and decisions are made only based on information about the past. Hence, the online model can be implemented in practice. At a certain moment it is decided whether a tour starts and which jobs it contains based on the number of all currently available jobs and their instantaneous processing time. This model can be implemented to refine the general batching strategy (i.e. average batch size and tour start intervals) that has been found in an optimization with the previously discussed offline models. If, for instance, the average batch size and the average length of tour start intervals are given, adjustments for exceptional busy days might be useful. The desired solution of the online model is to find the most efficient batch at one time step, i.e., a subset of all currently available jobs that is advantageous to achieve a good performance for the entire time horizon.

The objectives of an online approach can therefore not be the optimization of characteristics that evolve over time, such as the overall travel distance or average lead time. Rather, decisions are made that maximize performance with suitable performance measures at each time step.

To develop the model let first $t \in \mathcal{T}$ be a fixed time step. At this time step there is a set of pending jobs that arrived in the past and were not included in a previous tour. Let $\mathcal{J}^t \subset \mathcal{J}$ be this set of jobs. That is $\mathcal{J}^t = \{j_1, \ldots, j_F\}$, whereby for each $j_f \in \mathcal{J}^t$ the corresponding time component (cf. definition (1)) is equal to or less than *t*.

If employee *m* is available at time step *t*, the decision that has to be made first is whether a tour should start at time step *t* or not. There are several options to do that. We propose a decision based on the time gap between *t* and the arrival time of the earliest arrived job or the cardinal number of $\mathcal{J}^t$, respectively. In this case $r^{t,m}$ is defined as

$$r^{t,m} = \begin{cases} 1, & \text{there is a job } j = (n, t', s) \in \mathcal{J}^t\text{: } t - t' \geq t_{max} \\ 1, & |\mathcal{J}^t| \geq J^t_{max} \\ 0, & \text{otherwise} \end{cases} , \tag{22}$$

whereby $t_{max} \in \mathbb{R}_+$ is the maximal tolerated time gap between a t' and t and $J^t_{max} \in \mathbb{N}$ is the maximal number of pending jobs allowed.

The corresponding objective to the minimized overall walking distance at the tactical decision level for an online problem is the objective of maximizing an appropriate performance measure. We propose a performance measure that includes both tour length and batch size. Again, let $|J^{t,m}|$ be the number of products contained in a tour $J^{t,m} \subset \mathcal{J}^t$ of employee *m*. We define the performance measure

$$P^{\alpha}_{J^{t,m}} = \alpha \cdot |J^{t,m}| - L_{J^{t,m}} \qquad \forall J^{t,m} \subset \mathcal{J}^t \tag{23}$$

whereby $\alpha \in \mathbb{R}_+$ is used as a control parameter to weigh the importance of the batch size and the tour length against each other and also to control for differences in units. α can be chosen in practice by defining two performances of specific (different) solutions to be equal. In doing so, we determine how long a larger tour length is covered by a larger batch size in terms of the performance value. We remark that this performance measure is suitable for a comparison between batching performances with similar numbers of available jobs. If there are more elements in $\mathcal{J}^t$ efficient batches are easier to find than with fewer elements. So, performance is likely to be lower for small sets of available jobs. With the help of this measure we set the following objective for a time step $t \in \mathcal{T}$.

$$max\left\{P^{\alpha}_{J^{t,m}} \mid J^{t,m} \subset \mathcal{J}^t\right\} \tag{24}$$

Similarly a performance measure can be defined that weighs batch size against the average lead time of ordered products. When lead time is considered we have to take into account that lead times for products which are not contained in the following tour cannot be determined exactly. However, not included products must have an impact on the performance measure. Hence, a penalty value $c \in \mathbb{R}_+$ is added to the lead time of each not contained ordered product and we define one type of lead time for included products (25) and another type for not included products (26).

$$W_{j,1} = t - t' + D_{J^{t,m}} \qquad \forall j \in J^{t,m}, s = 0 \tag{25}$$
$$W_{j,2} = t - t' + D_{J^{t,m}} + c \qquad \forall j \in \mathcal{J}^t \backslash J^{t,m}, s = 0 \tag{26}$$

An estimation of the average lead time for all requested products is now given by

$$W_{ave} = \frac{1}{|\mathcal{J}^t_{ord}|}\left(\sum_{j_0 \in J^{t,m}} W_{j_0,1} + \sum_{j_0 \in \mathcal{J}^t \backslash J^{t,m}} W_{j_0,2}\right) \tag{27}$$

whereby $\mathcal{J}^t_{ord} \subset \mathcal{J}^t$ is the set of ordered jobs until t and $j_0 \in \mathcal{J}^t_{ord}$. A performance measure can now be defined by

$$P^{\beta}_{J^{t,m}} = \beta \cdot |J^{t,m}| - W_{ave} \qquad \forall J^{t,m} \subset \mathcal{J}^t \tag{28}$$

and as the resulting optimization target we obtain

$$max\left\{P^{\beta}_{J^{t,m}} \mid J^{t,m} \subset \mathcal{J}^t\right\}. \tag{29}$$

With this objective a solution algorithm tends to include the most urgent products in a batch and to form batches with many products. However, an incentive is created to merge products that are close to each other, in order to achieve short tour duration.

Similarly, a performance function to match the number of lateness's and batch size can be developed. Here, it should be taken into account that, if no lateness is going to occur, there is an incentive needed to reduce either tour length or waiting time.

Constraint (12) as well as the number of allowed late jobs (18) and the limitation of the batch size (19) can be taken over directly from the batching model at the tactical decision level, whereby their fulfillment is only required for time steps $t' \leq t$. Constraint (14) needs to be reformulated for one time step.

$$\sum_{m=1}^{M} inc_j^{t,m} = 1 \qquad \forall j \in \mathcal{J}^t \tag{30}$$

Moreover, the limitation of the lead time for ordered products (16) and for returned products (17) can be used for an online model, in which the objective is the maximization of labor efficiency as described in (24).

4. CATEGORIZATION OF MODELS

In the previous section we designed three optimization models to approach the batching problem for integrated forward and return flows. This section is dedicated to a categorization of these models concerning model and solution characteristics and an identification of possible solution opportunities based on methods that have been developed for the classical batching problem without return flow incorporation.

4.1 Model Characteristics

The deterministic as well as the stochastic model can be used to make tactical decisions. That is, they are designed to determine the general batching procedure, i.e., the average number of jobs in one batch and the average time between two tour starts, for a specific setting (warehouse layout, frequency of arriving jobs, and return rate). The online approach is designed to make decisions at the current time step. To identify a general batching procedure by using offline models might be insufficient to achieve a consistently high performance, due to daily fluctuations. Hence, after a general procedure is implemented the incorporation of the online model can help to refine the offline solutions by adjusting them to unforeseen demand or under-staffed days.

Regarding the problem class all models are highly complex. The deterministic and the online model are mixed integer nonlinear programs. The binary variables $inc_j^{t,m}$ and $r^{t,m}$ require integer solutions, whereas the length of a tour $L_{J^{t,m}}$ is a nonlinear function. The stochastic model belongs to the problem class of the nonlinear mixed integer multistage stochastic programs. Those problems are known to be high-dimensional and therefore difficult to solve.

Objective functions that can be used in the models can have either a labor efficiency or customer-oriented character. The deterministic model is designed to optimize objective functions that concern the overall time horizon, i.e.

performance measures such as the overall travel distance of employees or the average processing time of jobs are suitable for this model. The same applies to the stochastic model, whereby in this case the expected values of objective functions are supposed to be optimized. The online model is designed to facilitate solving the problem at a certain time step. Objective functions must therefore be defined based on available information. In section 3.3 we developed suitable performance measures for this model.

In all models different types of data are required to obtain solutions, since they are designed for different application. Historical data for varying setting classes is needed to derive solutions with the deterministic model. The stochastic approach requires an additional step of abstraction by determining the probabilistic characterization of the job flow and the generation of scenario trees. If the online approach is used, data is given by the actual current job flows.

The dimensionality of the model has already been addressed for the stochastic approach, which is highly dimensional due to the incorporated stochastic process. The deterministic model can become high dimensional when short time intervals and hence many time steps are considered, i.e. resolution in time needs to be high. The online approach on the other hand can be regarded as optimization problem with reduced dimension in comparison with the other two approaches, since performance is evaluated in only one time step.

4.1 Solution Characteristics

As mentioned previously, the application spectrum of the models is different. The deterministic and the stochastic model as can be used to identify batching policies that match the specific warehouse setting (cf. 3.1), i.e., a certain stock size, number of employees, frequency of job arrivals and so on. But those models can also be applied when vice versa the number of employees or the required transport capacity is supposed to be decided in order to achieve a desired performance, such as average processing time. The difficulty to model the stochastic process turns into an advantage when solutions are considered. Namely, the more precise the probabilistic characterization of the job flow is, the more consistent is the solution. Since decisions are only dependent on previous time steps and not on the entire time horizon as in the deterministic model, the solution reflects reality better than the deterministic solution. Moreover, the solution is dependent on the probabilistic characterization of the job flow, but independent from historical data (cf. 3.2). The online problem solution should provide the best way to react in a certain time step. The drawback of an online approach is that we cannot aim an overall optimal procedure along the entire time horizon. For that reason, the online approach should always be the second step after an offline solution has been identified and it should be used to refine this solution by adjusting the strategy based on current information (cf. 3.3).

To address the feasibility of the solutions their dependencies on controllable and uncontrollable factors need to be analyzed. The accuracy of the deterministic solution is highly dependent on the consistency of the setting class definitions, which is only partially controllable for the user. The classification of settings can be controlled, but if the job flow after the implementation differs considerably from the historical data that was used to find the alleged best batching policy of a certain setting class, it might be that the suitability of the solution is not assured. The scenario tree that is used in the stochastic model is supposed to reflect real job flows and should not be dependent on the historical data. With an online approach overall optimal solutions are not intended. However, the online solution is dependent on the control parameters of the underlying performance measure, which should be calibrated in a way that near-optimal solutions for the entire time horizon can be achieved.

The last aspect of solution characteristic that we consider is flexibility of the solution. Offline derived solutions are less flexible in terms of reaction to unforeseen events than online derived solutions, in which we react on the current and might be unexpected job flow.

To conclude this paper we focus on possible solution approaches for each of the models. There are several deterministic approaches to find solutions for the classical batching problem for forward flows. Albareda-Sambola et al. (2009) propose a heuristic method that is based on variable neighborhood search. Chen and Wu (2005) also use a deterministic approach, by presenting the association-based order-clustering model. Heuristics based on a cluster analysis for several underlying routing policies have been developed by Hwang and Him (2005). Optimal approaches are rare in literature due to the complexity of the problem already in the deterministic formulation. A near-optimal solution approach based on genetic algorithms has been developed by Hsu et al. (2005). A promising solution approach of the deterministic model for integrated forward and return flows could therefore be a heuristic based on genetic algorithms, as for forward flows proposed in Hsu et al. (2005).

Solution methods of the multistage stochastic model to solve the batching problem could not be found in literature. Methods that have been developed to solve multistage stochastic optimization programs for other application fields (financial mathematics and energy distribution systems) often use heuristic approaches, such as branch-and-bound algorithms, which might also be a good option to solve the batching problem for integrated forward and return flows.

An online approach for the classical batching problem has been proposed by Poon and Yu (2005) with the performance measure to describe the time needed to finish all jobs. Tian et al. (2007) present an online algorithm to minimize the time step in which all jobs are conducted. These methods can build the basis to design a solution algorithm for the online batching problem with integrated forward and return flows.

5. CONCLUSIONS AND FURTHER RESEARCH

The intention of this paper was to design models for batching with integrated forward and return flows and to clarify their application spectrum. A review of existing models to find batching policies for forward flows has shown that optimal offline solutions for integrated forward and return flows can most likely only be determined for small-sized problems. A focus on heuristic techniques (particularly to solve the online problem in reasonable time) is required. Therefore, we emphasize the need of solution methods for these problems. Moreover, it should be researched how the underlying routing policy, which is used to determine the tour length of a batch, affects the batching solution, or in other words, we need to analyze the stability of the solution concerning a change of the routing policy. Regarding routing policies another important issue to consider for incorporated returns is the batch size limitation, which might be treated more flexible. If the route of an employee is designed in a way, in which restoring activities are conducted first, the allowed batch size can be higher than the actual transport capacity. Particularly the performances of batching solutions that have been identified based on such routing policies are of high interest in this research field.

ACKNOWLEDGEMENTS
The authors are grateful to the generous support from Korea Research Foundation.

6. REFERENCES

Albareda-Sambola, M., Alonso-Ayuso, A., Molina, E., and de Blas, C. S. (2009). Variable neighborhood search for order batching in a warehouse. *Asia-Pacific Journal of Operational Research*, 26(5):655-683.

Bellman, S., Lohse, G. L., and Johnson, E. J. (1999). Predictors of online buying behavior. *Communications of the ACM*, 42(12).

Chen, M.-C. and Wu, H.-P. (2005). An association-based clustering approach to order batching considering customer demand patterns. *The International Journal of Management Science*, 33:333-343.

Consigli, G. and Dempster, M. (1998). Dynamic stochastic programming for asset liability management. *Annals of Operations Research*, 81:131-161.

De Koster, R., Le-Duc, T., and Roodbergen, K. J. (2007). Design and control of warehouse order picking: A literature review. *European Journal of Operational Research*, 182:481-501.

Gademann, N. and van de Velde, S. (2005). Order batching to minimize total travel time in a parallel-aisle warehouse. *IIE Transactions*, 37:63-75.

Gibson, D. R. and Sharp, G. P. (1992). Order batching procedures. *European Journal of Operational Research*, 58:57-67.

Heitsch, H. and Rőmisch, W. (2009). Scenario tree modeling for multistage stochastic programs. *Mathematical Programming*, 118:371-406.

Henn, S., Koch, S., Doerner, K., Strauss, C., and Wäscher, G. (2009). Metaheuristics for the order batching problem in manual order picking systems. FEMM Working Papers 09020, Otto-von-Guericke University Magdeburg, Faculty of Economics and Management.

Ho, Y.-C., Su, T.-S., and Shi, Z.-B. (2008). Order-batching methods for an order-picking warehouse with two cross aisles. *Computers and Industrial Engineering*, 55:321-347.

Hsu, C.-M., Chen, K.-Y., and Chen, M.-C. (2005). Batching orders in warehouses by minimizing travel distance with genetic algorithms. *Computers in Industry*, 56:169-178.

Hwang, H. and Him, D. G. (2005). Order-batching heuristics based on cluster analysis in a low-level. *International Journal of Production Research*, 43(17):3657-3670.

Petersen, C. G. (1997). An evaluation of order picking routing policies. *International Journal of Operations and Production Management*, 17(11/12):1098-1111.

Poon, C. K. and Yu, W. (2005). On-line scheduling algorithms for a batch machine with finite capacity. *Journal of Combinatorial Optimization*, 9:167-186.

Rouwenhorst, B., Reuter, B., Stockrahm, V., van Houtum, G., Mantel, R., and Zijm, W. (2000). Warehouse design and control: Framework and literature review. *European Journal of Operational Research*, 122:515-533.

Tian, J., Fu, R., and Yuan, J. (2007). On-line scheduling with delivery time on a single batch machine. *Theoretical Computer Science*, 374:49-57.

Won, J. and Olafsson, S. (2005). Joint order batching and order picking in warehouse operations. *International Journal of Production Research*, 43(7):1427-1442.

Wruck, S., Vis, I.F.A., Boter, J. (2010). Batching policies in Warehouses with High Return Flows (working paper VU University Amsterdam).

Session C2: Design & Operation 2

·**Day1: Sep. 15 (Wed.)**

·**Time: 15:00 - 16:20**

·**Chair: Amir Hossein Gharehgozli**

·**Room: Azalea & Lilac, 5F**

DESIGN METHODOLOGY FOR A NEW TYPE OF AUTOMATED CONTAINER TERMINAL

Lu Zhen, Loo Hay Lee*, and Ek Peng Chew

Department of Industrial & System Engineering
National University of Singapore,
10 Kent Ridge Crescent, 119260,
Singapore
e-mail: isezl@nus.edu.sg, iseleelh@nus.edu.sg, isecep@nus.edu.sg

Abstract: This paper introduces a new type of automated container terminal system, which utilizes frame bridges, rail-mounted frame trolleys and ground trolleys to transport containers between quay side and yard side. For this new design concept, this paper makes an explorative study to identify its challenges and opportunity. We investigate the proper resource requirements for implementing this new terminal system. A design methodology is proposed for determining proper numbers of resources in the system with given a design objective from port operators. This paper could help port operators who are interested in building new automated terminals, to better understand the relative merits of this new design and decide whether it is applicable in their terminals.

1. INTRODUCTION

The growth of global container shipping has trigged remarkable development and investment in ports. Until the recent economic crisis, which witnessed port throughputs fell by record numbers, there has been a wide consensus that the growing of demand for port handling capacity will outpace the development of the port infrastructure. Moreover, when the economic crisis is over and recovery starts, the strain on port handling capacity will likely be significant again. As a result, the concept of automated container terminal (ACT) has become very popular among port operators as a way to improve efficiency, reduce cost and increase capacity.

There are quite a number of different designs for ACTs, and many of them involve a combination of automated guided vehicles (AGV), automated stacking cranes (ASC), and automated straddle carriers (SC) (Murty *et al.*, 2005; Steenken *et al.*, 2004; Vis and de Koster, 2003; Vis, 2006). Some designs have shown performance improvement over conventional terminals. The Europe Container Terminal (ECT) in Rotterdam is a good example of the AGV-ASC combination. A design based on straddle carriers has been implemented in the Autostrad Terminal of Brisbane port. However, an ACT utilizing AGVs and ASCs would probably need large initial investment and maintenance cost, which reduces its attractiveness to port operators. The AGV-ASC combination does not always operate more efficiently due to the relatively low speed of the AGVs. On the other hand, a straddle carriers based ACT would have a lower space utilization rate which makes it less viable in larger, busier container terminals. The shortcomings of existing ACT designs have motivated the development of less expensive and more efficient new designs of ACTs. One of the new emerging concepts scraps the common usage of AGVs alike, and proposes a novel multi-storey solution by using rail-equipped frame bridges. This new ACT design concept is referred to as Frame Bridge based ACT (FB-ACT) in this paper.

1.1 FB-ACT: FRAME BRIDGE BASED AUTOMATED CONTAINER TERMINALS

As the name suggests, FB-ACT is a new type of ACT system based on multi-storey frame bridges and rails, on which electric trolleys transport containers between the quay and the yard. Figure.1 is an illustration (not to scale) for a FB-ACT.

As shown in Figure.1, the transport vehicles in FB-ACT are two types of trolleys: ground trolleys traveling in vertical directions on the rails; frame trolleys traveling in horizontal directions on the two storeys of frame bridges, in which there is one frame trolley in each storey of each row. The connection between these two trolleys is the transfer platform, which is a type of rail-mounted bridge crane and could move slowly on the rails of the highest story in frame bridges. During the operation process, the transfer platforms usually stay at a cross point of a ground rail and a row of frame bridge, and serve a dedicated block for a relative long time. The main function of the transfer platforms is to lift (or put down) containers between the ground rails and the rails in multi-storey frame bridges; in this process the containers are also rotated for 90 degrees by the transfer platforms. On the yard side, a vertical yard layout is used and each storage block is accompanied by ground rails with ground trolleys. During the unloading operation, containers are put down from quay cranes (QC) to the frame trolleys and then transported horizontally along the rails. At the transfer platforms, the containers are rotated for 90 degree and put down from the frame trolleys to the ground trolleys. The ground trolleys then transfer the containers along the vertical rails to the planned stacking location. The loading operation is performed similarly in the reverse order. As shown in Figure.1, the number of transfer platforms would better be equal to the number of blocks, and the number of frame trolleys. When all blocks are in operation simultaneously, one transfer platform along with one frame trolley could dedicate to a block.

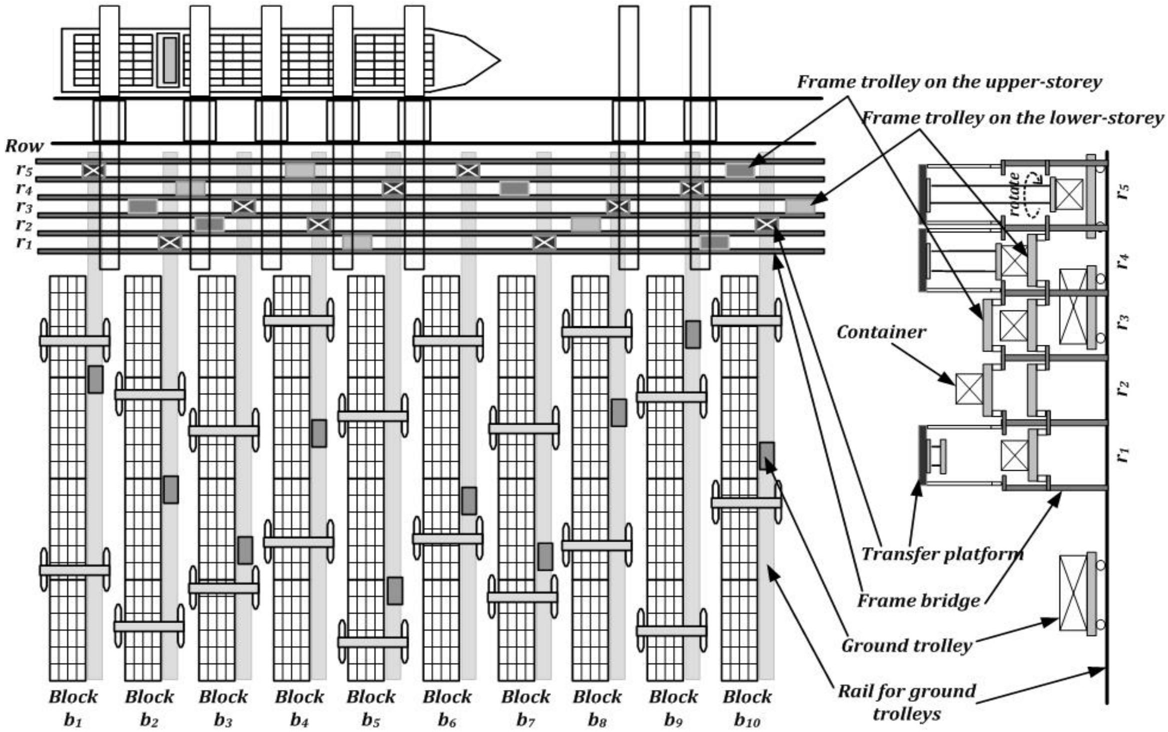


Figure.1 The configuration of the frame bridge based automated container terminal

Figure.1 illustrates an example of the FB-ACT system. In the realistic implementation, it could be extended in some aspects. For example, the frame bridges could be made with more than two storeys (as shown in Figure.2); the ground rails for ground trolleys could also be built with multi-storey structure. In this way, the handling capacity and efficiency of FB-ACT can be increased evidently and reach its full potential.

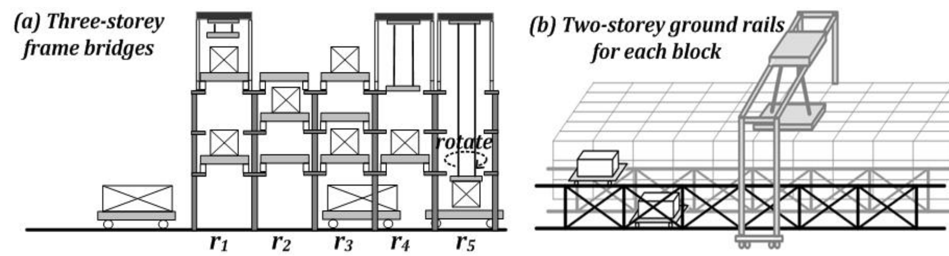


Figure.2 Some possible extensions for the FB-ACT system

This interesting new ACT design has several fundamental differences from other mainstream ACT designs. The first difference is that it removes the need for ground transportation vehicles (AGVs) and uses rail-mounted trolleys (flat cars), which reduces the initial investment as AGVs along with their controlling systems are typically quite expensive. In addition, the rail-mounted trolleys also promise higher speed than AGVs. Another merit is the elimination of excessive yard crane (YC) movement. In this design, the ground trolleys are able to deliver (or pick up) containers accurately to (from) the slot in blocks, which reduces the slow movement of YCs. These changes could increase the efficiency of the yard.

The advantages of FB-ACT are listed as follows: (1) pro-environment: it is powered by electricity without diesel engines, which is both cost effective and environmental friendly, (2) high stacking capacity: frame bridges and ground rails occupy small areas, (3) high efficiency: trolleys just run in their dedicated straight rails, thus they could run fast, and each trolley could carry two 40-ft containers, (4) simple control: there is no collision risk for trolleys, and no need of drivers, navigation equipments, nor complex control systems, and (5) low investments: it can also apply in the transformation of traditional terminal systems.

At present, FB-ACT has not been built and utilized by ports; but ZPMC (Shanghai Zhenhua Port Machinery Co. Ltd.) has implemented a real-size prototype system on an island of Shanghai in 2008. According to the news in *cargosystems.net*, this new ACT system will be built on the eastern side of Cao feidian port (the largest infrastructure project under construction in China). In addition, SSA Marine looks set to become the first terminal operator to install this ACT system in Long Beach within two years.

FB-ACT has sparked some challenges and opportunity for port operators. A lot of operation and planning mechanisms would require rethinking for this new ACT design. The existing researches mainly focus the analysis on the conventional or AGV based terminals. For example, Hoshino *et al.* proposed a methodology for evaluating vertical and horizontal AGV based ACTs (Hoshino *et al.*, 2007); Liu *et al.* utilized simulation method to compare four types of ACTs, e.g., AGV system, linear motor conveyance system, grid rail system, and automated storage and retrieval system (Liu *et al.*, 2002). Some analytical models were developed for designing AGV systems and studying their performances (Bozer and Srinivasan, 1991; Johnson and Brandeau, 1993). For the conventional terminals, Kim *et al.* recently proposed some analytical methodologies for optimizing the layout design (Kim *et al.*, 2008; Lee and Kim, 2010). To our best knowledge, few scholars have studied this new ACT system up to now. This paper makes an exploratory study on the design methodology for FB-ACT by investigating the resource requirements for satisfying the design objective.

2. DESIGN METHODOLOGY FOR FB-ACT

Suppose a port operator will build a FB-ACT with given a design target, e.g., the required throughput of terminal. Port designers should determine a set of proper numbers for: QCs (N_{QC}), frame trolleys (N_{FT}), rows (R), storeys (T) of frame bridges, transfer platforms (N_{TP}), blocks (N_{blk}), YCs (N_{YC}), ground trolleys for each block (N_{GT}), etc. Here we propose a design methodology for FB-ACT. More realistic factors will be taken into account, e.g., the different velocities of trolleys in laden or empty status. The aforementioned assumptions, e.g., full and even workload, one TP dedicates to one block, will be relaxed in the following study on design methodology for FB-ACT.

2.1 The framework of the design methodology for FB-ACT

The design objective is to determine the proper numbers of all resources in FB-ACT with the purpose of meeting the required throughput (π) for the terminal. With considering the fluctuation of workload in each block, we also need to know another design parameter, i.e., the times (η) of the peak throughput to the average throughput for one block. Figure.3 is the framework of methodology.

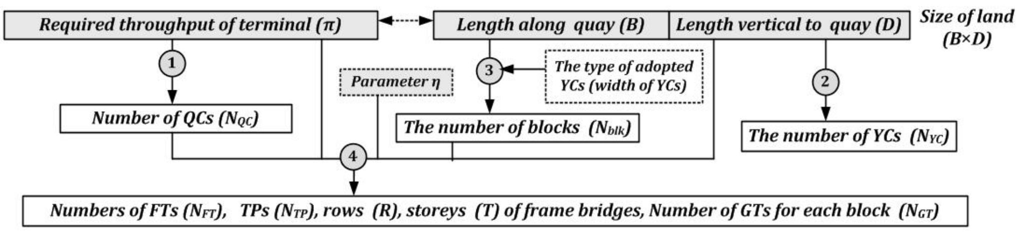


Figure.3 The framework of the design methodology for FB-ACT

As shown in Figure.3, the first step is to determine N_{QC} according to the required throughput π. The maximum rate for a QC is usually 40 tasks/h; hence we have $N_{QC} \geq \pi/40$. Then we decide N_{YC} on the basis of the land length perpendicular to the coastline (D). In the usual practice, two YCs are deployed in a block with length of 40 TEU slots. Here N_{YC} can also be determined by this rule accordingly. The third step is to determine N_{blk} based on the length of quay (B) and the width of each block that is dependent on the adopted type of YCs. Then, the fourth step determines the number of some key resources in FB-ACT.

Procedure for determining resource requirements // The 4th step in the design methodology

For N_{GT} = 1 **to** N_{GT}^{Max}, step 1 // N_{GT}^{Max}: the maximum number of GTs for a block; we set N_{GT}^{Max} = 2
  $w_{FT} = f(N_{GT})$; // Calculate the waiting time (w_{FT}) of FTs for GTs, according to N_{GT}
  For N_{FT} = 1 **to** N_{FT}^{Max}, step 1 // N_{FT}^{Max}: the maximum number of FTs; e.g., we set N_{FT}^{Max} as 4 times as N_{blk}
    $\pi' = T(N_{FT}, w_{FT})$; // Calculate the designed throughput (π') , according to N_{FT} and w_{FT}
    If $\pi' \geq \pi$ **Then** break; // If the designed throughput (π') satisfies the required throughput
  End For
  Determine R and T accordingly; // N_{FT} should be a multiple of R×T, i.e., N_{FT}%(R×T) = 0; R≤5; T≤3; this process may adjust (increase) N_{FT} slightly
  If $\pi' = T(N_{FT}, w_{FT}) \geq \pi$ **Then** break; // Reconfirm the designed throughput
End For

In the above procedure, $w_{FT} = f(N_{GT})$ and $\pi' = T(N_{FT}, w_{FT})$ are two key functions:

(1) $\boldsymbol{w_{FT} = f(N_{GT})}$

The number of ground trolleys for each block (N_{GT}) could be one; it may also be two if the length of block is very long. A proper N_{GT} should make the waiting time of frame trolleys (w_{FT}) be controlled around a reasonable value. Too short w_{FT} is not necessary, whereas too long w_{FT} will incur a bottleneck for the whole system. w_{FT} will be used in the next step for validating whether the designed throughput meets the required value. For how to calculate w_{FT} according to N_{GT}, i.e., $w_{FT} = f(N_{GT})$, the detailed calculation process is addressed in Appendix A.

(2) $\boldsymbol{\pi' = T(N_{FT}, w_{FT})}$

For analyzing the throughput, the system of frame trolleys is formulated as a closed cyclic queuing network model. A fixed number of frame trolleys circulate through the network at all times.

As shown in Figure.4, the closed queuing network contains four service facilities and N_{FT} circulating frame trolleys. Let the i^{th} facility has r_i parallel servers each with exponential distributed service time, the mean value of which is s_i. This assumption on the exponential distribution was also utilized in the similar studies on AGV systems (Hoshino *et al.*, 2007) or PM systems (Lee and Kim, 2010) by other scholars. As shown in Figure.4, $r_1 = N_{QC}$, $r_3 = N_{TP}$, $r_2 = r_4 = 15$; as we assume there are at most three storeys and five rows of frame bridges, 15 is the maximum number of rails in frame bridges.

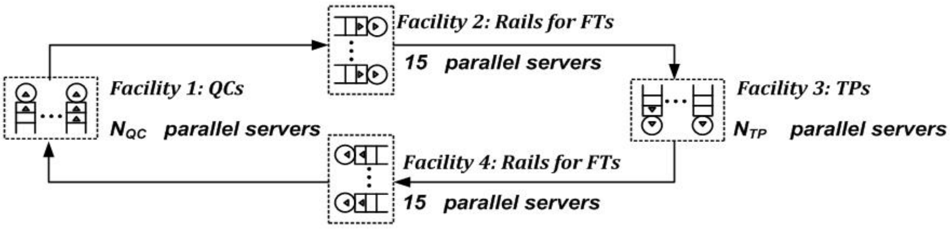


Figure.4 The closed cyclic queuing network model for frame trolleys

Note that the state of such a network can be described by a vector $\vec{n} = (n_1, \cdots, n_4)$ where n_i is the number of frame trolleys present at the i^{th} facility, $\sum_{i=1}^{4} n_i = N_{FT}$. According to Gordon and Newell (1967), the equilibrium distribution of frame trolleys in the network is given by:

$$Prob(n_1, \cdots, n_4) = \frac{1}{G(N_{FT})} \prod_{i=1}^{4} q_i(n_i) \quad (1)$$

$q_i(n_i)$ is the convolution parameter if n_i frame trolleys are at the i^{th} facility: (Hoshino *et al.*, 2007)

$$q_i(n_i) = \begin{cases} \dfrac{(\lambda_i \cdot s_i)^{n_i}}{n_i\,!}, & n_i \le r_i \\ \dfrac{(\lambda_i \cdot s_i)^{n_i}}{r_i\,! \cdot r_i^{(n_i - r_i)}}, & n_i > r_i \end{cases} \quad (2)$$

Here, λ_i is the arrival rate at the i^{th} facility. Since it is a closed cyclic network without branches, λ_i can be regarded as the arrival rate of loading/unloading tasks, i.e., π (tasks / h). s_i is the mean value of the 'service' time in the i^{th} facility. s_1, the waiting time of frame trolleys under QCs, is set with 1 min ; s_3, the waiting time of frame trolleys at transfer platforms, is obtained by $w_{FT} = f(N_{GT})$ in the previous step, $s_3 = w_{FT}$. For the travel time in laden and empty states, we calculate them as follows.

$$s_2 = \frac{B}{4v_{FT}^L}\left(1 + \frac{N^2-4}{3N^2}\right) \approx \frac{B}{3v_{FT}^L} \quad \text{// } v_{FT}^L\text{: the velocity of FT (frame trolleys) in laden state} \quad (3)$$

$$s_4 = \frac{B}{4v_{FT}^E}\left(1 + \frac{N^2-4}{3N^2}\right) \approx \frac{B}{3v_{FT}^E} \quad \text{// } v_{FT}^E\text{: the velocity of FT (frame trolleys) in empty state} \quad (4)$$

In Formula (1), $G(x)$ is the normalizing function when x frame trolleys are in the system. $G(x)$ is utilized to ensure all the possible $Prob(n_1, \cdots, n_4)$ sum to one (Gordon and Newell, 1967):

$$G(x) = \sum_{\vec{n} \in S(N,M)} \prod_{i=1}^{4} q_i(n_i) \quad (5)$$

Where $S(N, M) = \{(n_1, \cdots, n_4) \mid \sum_{i=1}^{4} n_i = x\}$.

Despite the apparently large number of arithmetic operations involved in $G(x)$, a simple algorithm was proposed to calculate it (Buzen, 1973), which is addressed in Appendix B.

Based on $G(x)$ function, the throughput of the closed cyclic queuing network is calculated as:

$$\pi' = \lambda_i \cdot G(N_{FT} - 1)/G(N_{FT}) \quad (6)$$

The above formula is actually the aforementioned function $\pi' = T(N_{FT}, w_{FT})$; and w_{FT} is already involved in the process for calculating $G(x)$. Here λ_i is the arrival rate of tasks at the i^{th} facility, which is also equal to π, i.e., the desired throughput for the terminal. Therefore, finding a proper N_{FT}^* to ensure the throughput satisfies the desired value is equivalent to:

$$N_{FT}^* = \min\{N_{FT} \mid G(N_{FT}) < G(N_{FT} - 1);\ N_{FT} \in \{1,2, \cdots, N_{FT}^{Max}\}\} \quad (7)$$

2.2 An example for FB-ACT design

An example is given to illustrate the process of the proposed design methodology. Some basic parameters are set as: v_{GT}^L, v_{FT}^L = 5.5 m/s, v_{GT}^E, v_{FT}^E = 6.5 m/s, N_{GT}^{Max} = 2, η = 2, h_{YC} = 1 min (for N_{GT} = 1), and h_{YC} = 1.5 min (for N_{GT} = 2). Given the available land with $B \times D$ = 420 × 400 m, the port operator wants to build a FB-ACT with the handling capacity (throughput) of 160 tasks/h.

Step1: As the maximum rate for a QC is usually 40 tasks/h, we have $N_{QC} \ge 4$. Hence we could deploy five QCs in the terminal.

Step2: The length of blocks: $L \approx D - 25 = 375$ m. In the usual practice of conventional terminals, two YCs are deployed in a block with length of 40 TEU slots (about 250 m). So we could deploy three YCs in each block.

Step3: If the type of YC with the span of 12 rows TEU (width of YC is about 35 m) is adopted for the terminal, the number of blocks N_{blk} = 420 / 35 = 12.

Step4: According to the procedure, we firstly set N_{GT} = 1, $w_{FT} = f(1)$ =1058 (too large). Then N_{GT} = 2, $w_{FT} = f(2)$ = 80 sec; thus we set N_{GT} = 2. For N_{FT} from 1 to 15, we calculate $\pi' = T(N_{FT}, 80)$ according to Formula (6). The ratio of actual throughput to the required value (π) is:

Table.1 The impact of N_{FT} to the throughput for the design example

N_{FT}	2	3	4	5	6	7	8	9	10	11	12	13	14	15
π'/π	0.24	0.36	0.48	0.60	0.72	0.84	0.96	1.08	1.20	1.32	1.44	1.55	1.67	1.78

From the above results, 9 is the least number of frame trolleys for satisfying the required throughput. Based on N_{FT}, we determine R (number of rows of the frame bridges) and T (number of storeys) accordingly; here N_{FT} should be a multiple of R×T, i.e., N_{FT}%(R×T) = 0; R≤5; T≤3. We could set R×T = 3×3 or 5×2, which may need adjust (increase) N_{FT} = 10. With considering lower frame bridges may be easier to built and incur lower construction costs, we could determine N_{FT} =10, and R×T = 5×2. For N_{TP}, i.e., the number of transfer platforms, we set N_{TP} =10 accordingly, which means two transfer platforms are deployed in each row of frame bridges.

2.3 Further analysis on the design example

By the proposed methodology, we can also analyze the performance of a FB-ACT design.

For a FB-ACT design with N_{FT} frame trolleys, the average number of frame trolleys at i^{th} facility, denoted by $\sigma_i(N_{FT})$, and the traffic intensity, denoted by $\rho_i(N_{FT})$, can be calculated as follows:

$$\sigma_i(N_{FT}) = \lambda_i \cdot s_i \cdot G(N_{FT}-1)/G(N_{FT}) \quad (8)$$

$$\rho_i(N_{FT}) = (\lambda_i \cdot s_i/r_i) \cdot G(N_{FT}-1)/G(N_{FT}) \quad (9)$$

Here r_i is the number of parallel servers in each facility. $r_1 = N_{QC}$= 5; $r_3 = N_{TR}$= 10; $r_2 = r_4$ = R×T = 10. For the design example in the previous section, the results of $\sigma_i(N_{FT})$ and $\rho_i(N_{FT})$ are shown in Table.2, from which we could see the bottleneck may locate at quay side.

Table.2 The average number of frame trolleys and traffic intensity at the four facilities

		QCs	Rails	TPs	Rails
Average Num. of FTs	$\sigma_i(10)$	3.2	1.2	4.2	1.4
Traffic Intensity	$\rho_i(10)$	64%	12%	42%	14%

3. CONCLUSIONS

This paper introduces a new type of ACT system, and makes an explorative study to identify its challenges and opportunity for it to be applied in container terminals. The contributions include: we investigate the proper resource requirements for implementing this new ACT system. A design methodology is proposed for determining proper numbers of resources in the system. Given a design objective (e.g., the required throughput), the port operators can estimate a proper combination of various resources (e.g., the required numbers of trolleys, rows and storeys of frame bridges, storeys of ground rails) by following our proposed design methodology. Then port operators reckon up the gross investment, which can help them assess the benefits and risks of this new ACT system.

However, there are still some limitations for this study: for the interest of simplicity, we simplify the consideration on the YCs, QCs, whose stochastic factors are not taken into account. The handling time (or waiting time) for these equipments is assumed to be constants in our study.

A comprehensive and realistic simulation study could remedy these defects of the analytical approaches proposed in this paper. In the future, we will develop a microscopic simulation test bed that allows detailed simulation and evaluation of trolley movements in the port area, along with container loading and unloading operations in an integrated manner. The simulation model should provide extensive measures of effectiveness for this new ACT system, and will be validated with real-world data from selected ports so as to objectively evaluate the performance of this new ACT system in the realistic environments.

ACKNOWLEDGEMENTS

The authors are most grateful to MPA (Maritime and Port Authority, Singapore) and NOL (Neptune Orient Lines) fellowship for financial supports that make this research possible.

4. APPENDICES

Appendix A. Calculation process for $w_{FT} = f(N_{GT})$

Notations: the number of ground trolleys for each block (N_{GT}), the required throughput for the terminal (π), the times of the peak throughput to the average throughput for one block (η), the number of blocks (N_{blk}), the vertical length of the terminal (D), the velocities of laden and empty ground trolleys (v_{GT}^L, v_{GT}^E), and the waiting time for YC's handling (h_{YC}), the expected waiting time of frame trolleys (w_{FT}).

We employ the theory of M/M/c queueing system to estimate the waiting time of frame trolleys. As the average throughput for each block is π/N_{blk}, the peak throughput for each block is: $\eta \cdot \pi/N_{blk}$. For each block, the N_{GT} ground trolleys could be regarded as N_{GT} parallel servers, and frame trolleys could be regarded as customers, whose arrival rate is $\eta \cdot \pi/N_{blk}$. The expected handling time for the 'server' is $E(T) = E(t_{GT}^L + t_{GT}^E + h_{YC})$. Here $t_{GT}^L = d_{GT}/v_{GT}^L$, $t_{GT}^E = d_{GT}/v_{GT}^E$, and d_{GT} is the random variable that denotes the distance between the pickup/delivery position in a block and a position in the rows of frame bridges. $E(d_{GT}) = D/2$, here D is the vertical length of the terminal. Then:

$$E(T) = E(d_{GT}/v_{GT}^L + d_{GT}/v_{GT}^E + h_{YC}) = \frac{(v_{GT}^L + v_{GT}^E)\cdot D}{2\cdot v_{GT}^L \cdot v_{GT}^E} + h_{YC} \quad (10)$$

According to the formula of the mean waiting time for the M/M/c queuing system, the waiting time of frame trolleys could be calculated as:

$$E(w_{FT}) = \begin{cases} \rho E(T)/(1-\rho), & if\ N_{GT} = 1 \\ \rho^2 E(T)/(1-\rho^2), & if\ N_{GT} = 2 \end{cases} \quad (11)$$

Here, $\rho = \eta \cdot \pi \cdot E(T)/(N_{blk} \cdot N_{GT})$.

Appendix B. Algorithm for calculating $G(x)$

For $x = 0$ to N_{GT}^{Max}, step 1 // Initialize $G(x)$
    If $x = 0$ Then $G(x) = 1$;
    Else $G(x) = 0$;
End for
For $i = 1$ to 4, step 1 // For all the four facilities
    For $x = 0$ to N_{GT}^{Max}, step 1 // For all possible number of frame trolleys

$$q_i(x) = \begin{cases} \frac{(\lambda_i \cdot s_i)^x}{x!}, & x \le r_i \\ \frac{(\lambda_i \cdot s_i)^x}{r_i! \cdot r_i^{(n_i - r_i)}}, & x > r_i \end{cases}$$ // Calculate the convolution parameter according to (2)

    End For
    For $x = N_{GT}^{Max}$ to 0, step -1

$$G(x) = \sum_{n=0}^{x} [q_i(n) \cdot G(x-n)]$$

    End For
End For

5. REFERENCES

Bozer, Y.A. and Srinivasan, M.M. (1991). Tandem configurations for automated guided vehicle systems and the analysis of single vehicle loops, *IIE Transactions* 23, 71-82.

Buzen, J.P. (1973). Computational algorithms for closed queuing networks with exponential servers, *Communications of the ACM* 16, 527-531.

Gordon, W.J. and Newell, G.F. (1967). Closed queuing systems with exponential servers, *Operations Research* 15, 254-265.

Hoshino, S., Ota, J., Shinozaki, A. and Hashimoto, H. (2007). Hybrid design methodology and cost-effectiveness evaluation of AGV Transportation systems, *IEEE Transactions on Automation Science and Engineering* 4, 360-372.

Johnson, M.E. and Brandeau, M.L. (1993). Analytic model for design of a multivehicle automated guided vehicle system, *Management Science* 39, 1477-1489.

Kim, K.H., Park, Y.M. and Jin, M.J. (2008). An optimal layout of container yards, *OR Spectrum* 30, 675-695.

Lee, B.K. and Kim, K.H. (2010). Optimizing the block size in container yards, *Transportation Research Part E* 46, 120-135.

Liu, C.I., Jula, H. and Ioannou, P.A. (2002). Design, simulation, and evaluation of automated container terminals, *IEEE Transactions on Intelligent Transportation Systems* 3, 12-26.

Liu, J.Y., Wan, Y.W. and Wang, L. (2006). Quay crane scheduling at container terminals to minimize the maximum relative tardiness of vessel departures, *Naval Research Logistics.* 53, 60-74.

Murty, K.G., Liu, J. Y., Wan, Y.W. and Linn, R. (2005). A decision support system for operations in a container terminal, *Decision Support Systems* 39, 309-332.

Petering, M.E.H. and Murty, K. G. (2009). Effect of block length and yard crane deployment systems on overall performance at a seaport container transshipment terminal, *Computer & Operations Research* 36, 1711-1725.

Petering, M.E.H. (2009). Effect of block width and storage yard layout on marine container terminal performance, *Transportation Research Part E* 45, 591-610.

Stahlbock, R. and Voß, S. (2008). Operations research at container terminals: a literature update, *OR Spectrum* 30, 1-52.

Steenken, D., Voß, S. and Stahlbock, R. (2004). Container terminal operation and operations research - A classification and literature review. *OR Spectrum* 26, 3-49.

Vis, I.F.A. and de Koster, R. (2003). Transshipment of containers at a container terminal: An overview, *European Journal of Operational Researches* 147, 1-16.

Vis, I.F.A. (2006). Survey of research in the design and control of automated guided vehicle systems, *European Journal of Operational Researches* 170, 677-709.

SEQUENCING CONTAINER STORAGES AND RETRIEVALS

AH Gharehgozli[1*], Yugang Yu[1], R de Koster[1], JT Udding[2]

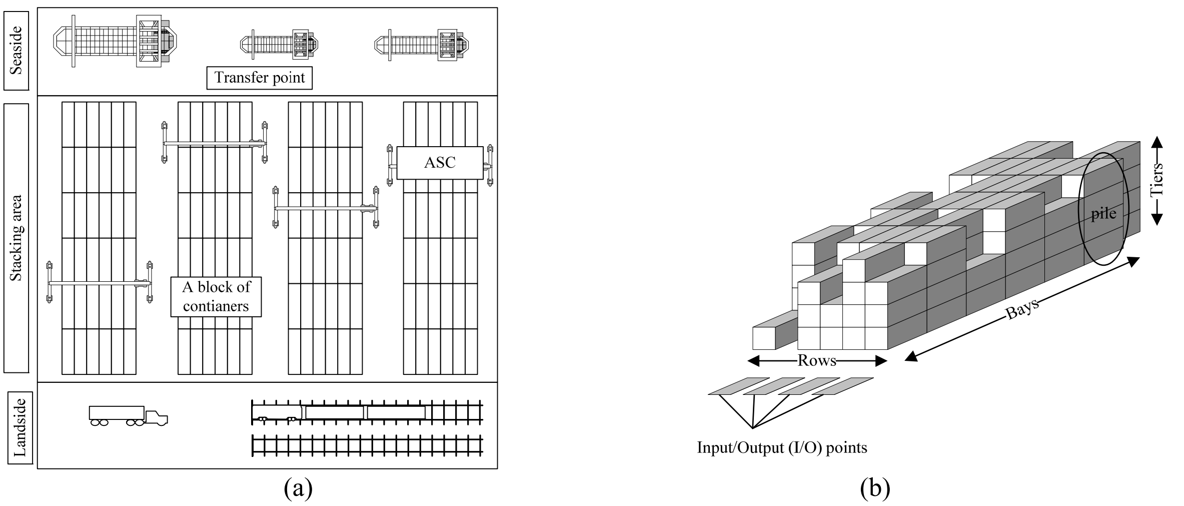
[1]Rotterdam School of Management
Erasmus University Rotterdam
PO BOX 1738, 3000 DR
Rotterdam, Netherlands
e-mail: {AGharehgozli, YYugang, RKoster}@rsm.nl

[2]Mechanical Engineering, Systems Engineering
Eindhoven University of Technology
PO Box 513, WH 0.129, 5600 MB
Eindhoven, Netherlands
e-mail: J.T.Udding@tue.nl

1. INTRODUCTION

Figure 1-a shows a typical layout of a container yard with several blocks in the stack area where automated stacking cranes (ASCs) store and retrieve containers. Each crane can handle containers in one block of the stack consisting of multiple rows, tiers, and bays as shown in
Figure 1-b. Within one block, containers are densely stacked in piles. A number of input/output (I/O) points are located at the seaside and landside of the block.

The ASC can move along the bays and rows of the block simultaneously. We can divide the movements of an ASC into two modes: single command and dual command cycles. In a single command cycle, the ASC starts at the I/O point, either stores or retrieves a container, and returns to the I/O point. In a dual command cycle, the ASC first picks up a container from an I/O point, travels to a stack location to store it, travels to another location to retrieve a container, and then returns to an I/O point to deliver it. In a dynamic environment, storage and retrieval requests randomly arrive at the block one after another. The basic sequencing strategy for an ASC is to perform requests in first-come-first-served (FCFS) order.



(a) (b)

Figure 1. (a) A container yard layout, (b) a block of containers

We consider the problem of sequencing storage and retrieval requests in a single block of containers given that the ASC works in dual command cycle mode. We assume a given number of M containers have to be stored and $N-M$ containers have to be retrieved from the block. Storage container has to be picked up from an I/O point and brought to given storage location. Retrieval container has to be brought to either a seaside or a landside I/O point. All moves are non-preemptive. Containers cannot be dropped off in a temporary location as this is time-consuming. The objective function is to minimize the makespan of the ASC to perform all requests.

The literature on this problem can be reviewed from two perspectives. Some papers directly discuss sequencing storage and retrieval requests in the storage area of a container yard or a warehouse. On the other hand, sequencing storage and retrieval requests is common in vehicle routing problems with unit loads (VRPUL). Next, we briefly discuss both streams of study.

1.1 Literature on sequencing storage and retrieval requests

Vis and Roodbergen (2009) consider scheduling storage and retrieval requests in a block of containers. They propose an exact polynomial solution method based on dynamic programming. However, they assume container yard uses straddle carriers (SCs) to store and retrieve containers. This differs from our problem because when a SC enters a row of containers, all requests have to be finished before it leaves that row, since changing rows is time-consuming for a SC.

Vis and Carlo (2010) propose a model to sequence storage and retrieval requests in a block of containers with two ASCs. The ASCs can pass each other but cannot perform operation on the same bay simultaneously. They propose simulated annealing to solve this problem. In this problem, instead of multiple I/O points only one I/O point exists at each side of the block.

Van den Berg and Gademann (1999) consider sequencing storages and retrievals in an automated storage/retrieval system (AS/RS). In an AS/RS system, an automated storage/retrieval (S/R) machine stores and retrieves unit loads. They formulate the problem as a transportation problem and claim each feasible solution of the transportation problem corresponds to a retrieval and storage sequence of the main problem. They assume a single commend input point and a single commend output point. Furthermore, the problem is 2 dimensional.

Zhang et al. (2010) work on a similar problem in an AS/RS system with two I/O points. However, retrievals and storages can be performed in both I/O points. They propose an exact polynomial solution method. The method is based on the assignment problem (AP) relaxation and a subtour patching scheme.

1.2 Literature on vehicle routing problem with unit loads

Our problem can be formulated as both the asymmetric travelling salesman problem (ATSP) and the Stacker crane problem (SCP). Note that our problem has special properties that make it different than all the following studies. In our problem, every request is connected to an I/O point and retrieval requests have flexible I/O points.

The ATSP, most simply stated, is a combinatorial problem in which from a given list of cities with pairwise distances, a tour has to be determined passing every city exactly once in such a way that the total distance is minimized. Pairwise distances need not be equal. Several algorithms are proposed to solve the ATSP. Among the most efficient ones is the depth-first branch and bound algorithm based on the subtour elimination approach proposed by Carpaneto and Toth (1980). In general, the ATSP problem is $\boldsymbol{NP}$-hard.

The SCP, in its simplest form, can be described as follows: a vehicle must start from an initial location, perform a specified set of moves, and return to its initial location. The moves are defined as carrying moves from a specified pickup point to a delivery point. The objective is to serve all required carrying moves with the shortest empty distance. This is done by adding non-carrying (empty) moves.

The SCP was originally proposed by Fredrickson et al. (1978). They prove that the SCP is $\boldsymbol{NP}$-hard when the underlying network is a complete graph. Fredrickson and Guan (1993) prove that the SCP on a tree when the movements are non-preemptive is also $\boldsymbol{NP}$-hard. They prove it by reducing the Steiner tree problem in bipartite graphs to the non-preemptive SCP.

2. PROBLEM DESCRIPTION

Our objective is to minimize the total makespan of the ASC for a group of storage and retrieval containers in a block with multiple I/O points. This problem can be formulated as the ATSP. The model consists of an AP model plus a subtour elimination constraint.

Consider that V is the set of all locations: storages, retrievals, ending point (mentioned as 0) and starting point (mentioned as 0'). Location of a storage container is the location of the bay that it has to be delivered and location of a retrieval container is the location of the bay that it is currently stacked. The crane can start from any bay and finish in any bay. The distance matrix, D, is a $(N+2)\times(N+2)$ matrix where, N is the number of storage and retrieval requests and 2 is the number of ending and starting points. Based on the set of locations, V, and the set of arcs, A, a complete digraph $G(V, A)$, can be defined. The associated distance d_{ij} of each arc $(i, j) \in A$, comes from the distance matrix D. Another graph $G'(V', A')$, can be also defined, where $V' = \{v_1, v_2, \dots, v_p\} \subseteq V$ is a subset of locations with corresponding arcs as $A' = \{(v_1, v_2), (v_2, v_3), \dots, (v_p, v_1)\} \subseteq A$. If $p = N$, the ASC traverses a Hamiltonian tour; otherwise ($p \leq N$), a Hamiltonian subtour. The makespan of the ASC traversing all arcs of a tour (subtour) equals the sum of distances of all arcs.

In this setting, the objective function is to find a Hamiltonian tour with the minimum makespan given the complete digraph G. The ATSP corresponding to the sequencing of storage and retrieval requests may be mathematically stated as the following integer programming model.

$$\min f(x) = \sum_{i \in V \setminus \{j, 0'\}} \sum_{j \in V \setminus \{0\}} d_{ij} x_{ij}, \quad (1)$$

Each location j must be entered exactly once. We cannot enter the starting point from any other location:

$$\sum_{i \in V \setminus \{j\}} x_{ij} = 1, \ \forall j \in V \setminus \{0\}, \quad (2)$$

Each location i must be exited exactly once. We cannot exit from the ending point to any other location:

$$\sum_{j \in V \setminus \{i\}} x_{ij} = 1, \ \forall j \in V \setminus \{0'\}, \quad (3)$$

No subtours are allowed for any subset of locations $S \subset V \setminus \{0', 0\}, S \neq \emptyset$, where $|S|$ is the size of S:

$$\sum_{i,j \in S} x_{ij} \leq |S| - 1, \quad (4)$$

All the variables are binary:

$$x_{ij} \in \{0,1\}, \ \forall i, j \in V \quad (5)$$

Equations (1)- (3) and (5) define the AP problem. The number of feasible solution of the AP model is N! for N locations of which $(N-1)$! are feasible Hamiltonian tours for the ATSP. A feasible solution of the AP model is a feasible solution of the ATSP if it satisfies equation (4), the subtour elimination constraint. The ATSP model is more constrained than the AP model hence, the optimal value of the AP is a lower bound for the ATSP.

The ATSP model is in general ***NP***-hard. However, the computation time of the problem of sequencing storage and retrieval requests mainly depends on the number of I/O points. For few number of I/O points, exact polynomial solution methods can be developed. In case of 2 I/O points, Zhang et al. (2010) have already developed a solution method. Their solution method can be extended to more number of I/O points using enumeration. Nevertheless, the computation time is not reasonable. This is due to the fact that with the increase of the number of I/O points, the computation time of the problem increases exponentially. This argument is supported by reducing the stacker crane problem (SCP) with non-preemptive movements on a general graph to the problem of sequencing storage and retrieval requests.

Consider our problem in a simplified case with the following assumptions: (1) every retrieval request has to be delivered to a specific I/O point, (2) the number of storage and retrieval requests is equal to the number of I/O points, and (3) every I/O point is connected to only one retrieval or storage request. In this setting, we have a problem with some carrying moves and we need to add non-carrying moves in order to get a complete tour with the shortest empty distance. This implies that the problem is an SCP that is ***NP***-hard (Fredrickson and Guan, 1993).

3. SOLUTION METHOD

We extend the branch and bound (B&B) algorithm proposed by Carpaneto and Toth (1980) for the ATSP to this problem. We also benefit from some of the ideas mentioned by Miller and Pekny (1991). The B&B algorithm proposed in this paper mainly differs from the previous ones regarding the following aspects:

(1) We use the depth first search as the overall strategy of the B&B algorithm to select the next subproblem and the best first search when choice is to be made between nodes at the same level.
(2) We produce a spars distance matrix based on the characteristics of the problem. The sparse distance matrix helps us to avoid explicitly examining a vast majority of feasible solutions.

(3) The other important issue is to select a subtour to be split. A smart method for selecting a subtour results in fewer number of nodes to be evaluated. We propose to select the subtour which contains the least number of I/O points.
(4) Furthermore, we add an extra step to the algorithm. In each node, we evaluate if it is possible to merge some of the subtours using the method proposed by Zhang et al. (2010). This also results in evaluating less number of nodes.

4. CONCLUSION

We showed that the computation time of the problem increases exponentially when the number of I/O points increases. However, if the number of I/O points is limited, an exact polynomial solution method can be developed. We propose a B&B algorithm to get the optimal solution of the model. The proposed B&B algorithm differ from previous ones proposed for solving ATSP in several aspects.

5. REFERENCES

Carpeneto, G., Toth, P. (1980). Some New Branching and Bounding Criteria for the Asymmetric Travelling Salesman Problem. *Management Science*, 26(7): 736-743.

Frederickson, G.N., Guan, D.J. (1993). Nonpreemptive ensemble motion planning on a tree. *Journal of Algorithms*, 15(1): 29-60.t

Frederickson, G.N., Hecht, M.S., Kim, C.E. (1978). Approximation Algorithms for Some Routing Problems. *ACM Journal of Computing*, 7(2): 178-193.

Miller, D.L., Pekny, J.F. (1991). Exact Solution of Large Asymmetric Traveling Salesman Problems. *Science*, 251(4995): 754-761.

van den Berg, J.P., Gademann A.J.R.M. (1999). Optimal routing in an automated storage/retrieval system with dedicated storage. *IIE Transactions*, 31(5): 407-415.

Vis, I.F.A., Carlo H.J. (2010). Sequencing Two Cooperating Automated Stacking Cranes in a Container Terminal. *Transportation Science*, 44(2): 169-182.

Vis, I.F.A., Roodbergen K.J. (2009). Scheduling of Container Storage and Retrieval. *Operations Research*, 57(2): 456-467.

Zhang, X., Yu, Y., van de Velde, S., de Koster, R. (2010). Minimizing the makespan of storage and retrievals in a two-depot multi-aisle AS/RS: a solvable ATSP problem. *working paper*.

AN ITERATIVE SEARCH ALGORITHM FOR LOAD SCHEDULING AND STORAGE ALLOCATION IN A MARITIME CONTAINER TERMINAL

Lu Chen[1*] and André Langevin[2]

[1]School of Mechanical Engineering
Shanghai Jiao Tong University
200240
Shanghai, China
e-mail: chenlu@sjtu.edu.cn

[2]Département de Mathématiques et de Génie Industriel,
École Polytechnique de Montréal and CIRRELT
H3C 3A7
Canada
e-mail: andre.langevin@polymtl.ca

Abstract: The load scheduling and the storage allocation for outbound containers are two important decision problems affecting the efficiency of container loading operations. This paper proposes a novel approach that integrates these two problems into a whole. The objective is to minimize the completion time of the loading operations. Due to the intractability of the proposed problem, an iterative search algorithm is designed to solve the problem. A decision is made upon the arrival of each container at the terminal. Computational experiments are conducted to examine the key factors of the problem and the performance of the proposed algorithm.

1. INTRODUCTION

The globalization of international markets has significantly fuelled the growth of the containerized freight transport industry in the last several decades. The overall performance of a container terminal is usually measured by the ship discharging/loading time, but the efficiency can only be achieved if the planning and control functions at different handling stages of the container flow are properly coordinated.

Generally speaking, a container terminal consists of the quayside, where containers are discharged from and loaded onto ships, and the container storage yard, where containers are stored. In the loading operations, the outbound containers are picked up by Yard Cranes (YCs) from the storage yard, transported from the storage yard to the quayside by trucks, and loaded onto the ship by Quay Cranes (QCs). The storage yard in the terminal is divided into the several blocks of containers. Each block consists of a number of side by side lanes with each lane including a number of container stacks that are of 4-5 tiers of containers.

During the loading operation, various constraints pre-specified by carriers must be satisfied. Examples are weight constraints, destinations and sizes of containers loaded into specific slots of the container ship. Thus, the loading schedule has to be carefully constructed before the loading operations begin. Also, the temporary storage of the outbound containers is one of the most important operations at the container terminal that has an important impact on the handling efficiency on the yard crane scheduling.

This paper focuses on the load scheduling and storage allocation problems (LS-SAP) for outbound containers, with the objective to enhance the efficiency of the loading operations in container terminals. Followed by this introductory section and a literature review in Section 2, the problem is defined in Section 3. In Section 4, an iterative search algorithm is developed to solve the problem. Computational experiments are reported in Section 5 and the performance of the algorithm is analyzed. Finally, Section 6 concludes this paper.

2. LITERATURE REVIEW

Problems associated with allocating and scheduling the resources in a container terminal have been extensively studied in the past few years. See Vis and De Koster (2003), Steenken et al. (2004), and Stahlbock and Voß (2008) for comprehensive reviews of the literature on various models and aspects of the optimization problems in a container terminal. Here, we provide a brief review of existing studies on YC scheduling and storage allocation.

In the area of load scheduling, Kim and Kim (2002) developed a cost model for determining the optimal amount of storage space and the optimal number of transfer cranes for handling inbound containers. Kim and Kim (1999a,b,c) tried to minimize the total travel distance of a straddle crane in the yard during the loading operation of outbound containers. The solution procedure is decomposed into two stages. In the first stage, the number of containers to pick up at each yard bay is determined, while the visiting sequence of yard bays by the straddle crane is determined in the second stage. Jung and Kim (2006) extended the scheduling problem for a single YC to multiple YCs within a block. Non-crossing constraints were considered. The objective is to minimize the makespan of the loading operations. A genetic algorithm and a simulated annealing algorithm are developed to solve the problem. Ng and Mak (2005) studied the problem of scheduling a single YC. The loading/unloading containers to be handled are considered as a given set of jobs with different ready times. The objective is to minimize the sum of job waiting time. The authors proposed a branch-and-bound algorithm to solve the problem optimally. Ng (2005) extended the study to multiple YCs. To solve the problem, a two-phase heuristics is developed. In the first phase, the multiple YCs scheduling problem is decomposed into independent single YC scheduling problem. Then, in the second phase a container re-assignment procedure is applied to improve the schedule.

Lee et al. (2007) studied the scheduling problem of two YCs which serve one QC at two different container blocks. The objective is to minimize the total transferring time in the stack area. A simulated annealing algorithm is developed to determine both the visiting sequences of yard bays and the number of containers picked up at each bay for the two YCs. Kozan and Preston (1999) used genetic algorithm techniques to reduce container handling time at a container terminal, and analyzed the major factors that influence container transfer efficiency. Effects of the handling equipment, the storage capacities and policies, and the terminal layout are analyzed through simulation. Kozan and Preston (2006) extended the work by integrating the container storage assignment policy with the scheduling of the handling equipment. The aim of the model is to simultaneously optimize both the storage locations and the handling scheduling.

For storage allocation problems, Dekker et al. (2006) explored different stacking policies for containers in automated terminals by means of simulation. A comprehensive overview of stacking policies used in practice is provided. Zhang et al. (2003) modeled the storage allocation problem using a rolling-horizon approach. For each planning horizon, the problem is decomposed into two levels, and each level is formulated as a mathematical programming model.

Preston and Kozan (2001) used genetic algorithm technique to determine the way that outbound containers are stored to minimize the time spent transferring containers from a storage area to ship or ship to the storage area. Kim and Park (2003) decomposed the process of determining the storage locations for outbound containers into two stages: space allocation stage and stage of locating individual containers. The authors developed a mixed integer program to solve the space allocation problem in the first stage, with the objectives to minimize the containers' delivery cost between the berth and the yard and to minimize the traveling cost of YCs to pick up the outbound containers within a certain range of yard area. Kim et al. (2000) focused on the stage of locating individual container by determining the storage location of an outbound container in a pre-assigned yard bay in order to reduce re-handle during loading operations. Dynamic programming is used to solve the problem.

Both load scheduling and storage allocation problems have attracted a lot of research efforts. However, these problems are highly related to each other. Kozan and Preston (2006) is the only study found so far that combines these two sub-problems. Tabu search and genetic algorithm are applied to simultaneously optimize both the storage locations and the load scheduling. However, the approach modeled each sub-problem separately. And some practical constraints are not considered, for example, the random arrival of outbound containers, and the weight and destination information of outbound containers, etc.

3. PROBLEM DEFINITION

Usually, outbound containers start arriving at the terminal waiting for loading two to three days before the ship departs. In the loading operation, containers to be loaded into slots of a ship must satisfy various constraints specified by the load file. A load file is made by the shipping line based on container groups, and is sent to the terminal. Before the arrival of the ship, a more detailed storage plan is made by the terminal planner who fills the groups with detailed containers. A container group is defined as a collection of containers that have the same characteristics such as the destination port, the weight class and the size.

Given the load file, this paper attempts to determine the loading sequence and the storage allocation for outbound containers at the same time. A loading sequence must be constructed to deliver the containers to the QCs satisfying the constraints specified by the load file regardless of locations of the containers to be loaded. In order to have an efficient loading sequence, outbound containers must be laid out in the optimal locations.

The rationale behind integrating the load scheduling and storage allocation is that it is pointless optimizing one when the other is far from optimal. The aim of the integrated model is to simultaneously optimize both the loading sequence and the storage locations for outbound containers.

The following assumptions and problem settings are imposed in defining the LS-SAP:

(1) We optimize the loading sequence, and don't consider the handling order of each YC. However, the detail YCs handling orders can be easily constructed given the loading sequence.

(2) There exist precedence constraints among different container groups in the loading operations. The load file of the corresponding ship, which indicates the storage slots of container groups on the ship and the number of containers in each group, is known.

(3) The handling time of each container in the yard is determined by the storage allocation plan, which means that it is a decision variable.

(4) Rehandle is taken into account in the model. A rehandle is a container movement made in order to permit access to another container, and is considered as an unproductive move.

The following decisions need to be made so as to minimize the makespan of handling the loading operations in a planning horizon: (i) determining the loading sequence for all the outbound containers; (ii) determining the storage location in the yard for each outbound container.

The ideal layout of outbound containers in the storage yard is almost impossible to be achieved due to the random arrival of containers at the terminal. We try to deal with this property in the development of the solution technique.

4. SOLUTION TECHNIQUE

Two heuristic algorithms are developed firstly to solve the two sub-problems independently. Then, the two algorithms are incorporated into a whole in an iterative way: the output of one algorithm is used as the input of the other, capturing the interactions between the two problems. The solution technique is explained as follows: a procedure based on the Greedy Randomized Adaptive Search Procedure (GRASP) is developed to determine the storage locations for outbound containers given the containers information, the container block configuration, and a loading sequence. Then, a Tabu Search algorithm is designed to schedule the loading operations given the storage locations specified in the previous step, the load file and the QCs work strategy. The optimized loading sequence is then used to re-allocate storage locations for the containers. The tabu search will re-start using the updated storage allocation plan. This procedure continues iteratively until a stopping criterion is reached. The framework of the iterative search is illustrated in Figure 1. This technique gradually increases the accuracy of the solution as the algorithm progresses, and gives a better overall solution.

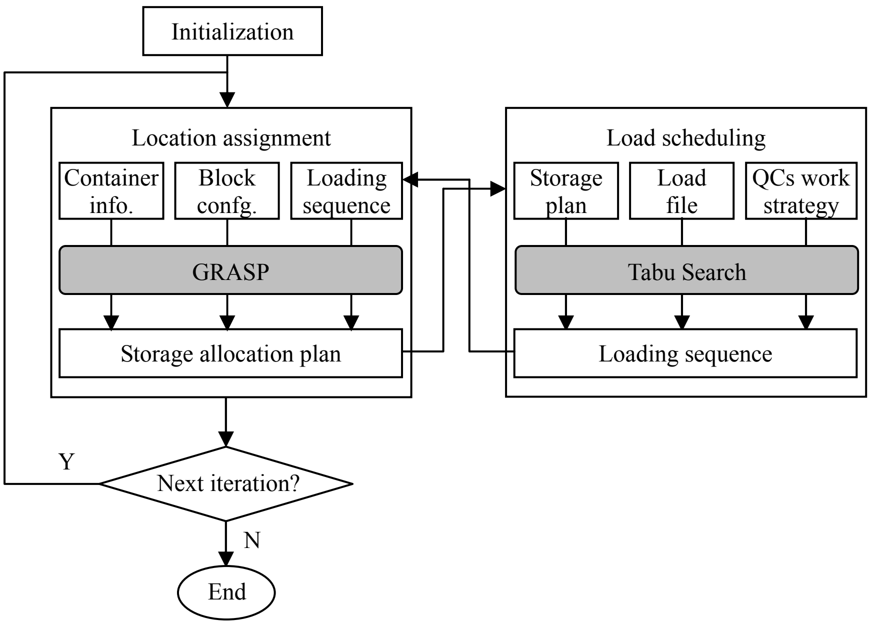


Figure 1. Framework of the iterative search algorithm

The follow notations are used in the introduction of the heuristic algorithm.

N: set of containers to be handled in the planning horizon.

K: set of storage locations for outbound containers in the storage yard.

p_i: processing time of container i ($i \in N$).

The handling time of container i, p_i, is the sum of the retrieval time and the rehandle time. If the desired container is on top there is no rehandle time, otherwise the rehandle time incorporates the time required to move those containers on top of the desired containers to the adjoining positions and return the top containers from the temporary storage positions back to the storage stack.

4.1 Storage allocation for outbound containers

Assume that a container block with empty slots is already assigned to a specific container group by a space planner. The remaining problems for determining the storage location are: (1) select a yard bay with empty slots in the pre-assigned block; (2) stack the container of a certain group among several empty slots in the selected yard bay. Since containers arrive at the terminal in a completely random order, this decision has to be made upon the arrival of each container. In addition, determining the storage location for outbound containers is a dynamic decision making problem due to the fact that the configuration of the yard bay changes each time a container is assigned to a location.

An initial loading sequence is assumed to be generated that is compatible with the work schedule of the QCs. In this section, a GRASP based heuristics is presented to determine the storage locations for outbound containers in the storage yard, respecting the handling sequence of the loading operations. The objective is to minimize the expected number of rehandle during the loading operations.

GRASP mainly consists of two phases: a construction phase and an improvement phase. In our procedure, the construction phase constructs a feasible solution using a biased sampling scheme. The feasible solution obtained in the construction phase is then improved by a 2-opt local search.

N is the set of the containers that need to be handled. Let $n = |N|$. Then there are n decision epochs. Let N^e denote the set of containers that have not arrived yet at epoch e, $e = 1, 2, \ldots, n$. Initially, $N^1 = N$. At each epoch, GRASP is applied to determine the storage locations for the containers that will arrive in the following epochs, taking into account the handling sequence of those containers, keeping the storage plan for the containers that have already arrived in the previous epochs. In order to describe the specific procedure of the GRASP, the index value of a container is calculated.

4.1.1 <u>The index values of containers</u>

The containers index values are calculated in order to take into account the handling sequence of the outbound containers when we try to determine their storage locations. It is expected that those containers to be retrieved earlier will be stacked in upper tiers in the storage yard. At epoch e, suppose container i is the jth one to be retrieved in the loading sequence in N^e, the index value of container i, V_i, is defined as follows:

$$V_i = |N^e| - j,\ i \in N^e$$

With the above definition, it's easy to see that those containers to be retrieved earlier will have bigger index value.

4.1.2 <u>The procedure of GRASP</u>

The GRASP construction procedure (GRASP-C) is described as follows.

Step 1: *Get index value level:* calculate the index value V_i for each unassigned container i, get the index value level l_i of each container.

Step 2: *Select a container:* select a container (denoted as container c) from those who have the smallest index values as the next one to be stacked.

Step 3: *Select a yard bay:* select in the block pre-assigned a yard bay that is not full and only with containers of the same destination with the selected container c.

Step 4: Stack container c in an available empty location (an empty slot is called "available" when all the slots below the empty slot are occupied) in the selected yard bay.

Step 5: Repeat Step 1 to Step 4 until all the containers in N^e have been assigned to locations. A feasible solution is generated.

The GRASP improvement procedure (GRASP-I) is described as follows:

Step 1. Calculate the objective value (the expected number of rehandle of the given loading sequence) of the feasible solution obtained previously and set it as the initial objective value.

Step 2. Set $x = 1$, $x' = x + 1$.

Step 3. Swap the locations of x^{th} and x'^{th} containers in the given loading sequence. Calculate the objective value of the resulting solution. If it is less than the initial objective value, update the initial value; otherwise, swap back the locations of the two containers as they were.

Step 4. Increase x' by 1, and go to Step 3 until $x' > |N|$.

Step 5. Increase x by 1, and go to Step 3 until $x > |N| - 1$.

The stopping criterion of GRASP is to repeat GRASP-C and GRASP-I n_{GRASP} times to generate n_{GRASP} solutions, where n_{GRASP} represents the iteration times. The solution with the least rehandle operations among the n_{GRASP} solutions is selected as the final solution and its corresponding objective value is set as the final objective value.

4.1.3 Stack the next coming container

The GRASP solution provides a near optimal storage location plan for the containers that will arrive in the following epochs. However, outbound containers arrive at the terminal in a completely random order. It may happen that storage location for the next arriving container is not available. Therefore, the following rule is applied when determining the storage location of the next arriving container, denoted as c.

If the storage location in the GRASP solution for c, denoted as k, is available, then stack c in k; otherwise, stack c in the available location that has the minimum distance from k.

The distance between two locations is calculated by putting a yard bay into a coordinate system as shown in Figure 2. Each slot represents a point in the coordinate system. The distance between location k and k' is their rectilinear distance: $d_{kk'} = |x_k - x_{k'}| + |y_k - y_{k'}|$.

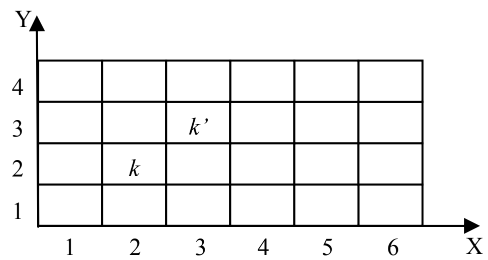


Figure 2. An illustration of a bay coordinate system

4.2 Tabu search algorithm for scheduling the loading operations

Given the load file of the ship and the locations of the outbound containers, a loading sequence of the YCs must be constructed to deliver the containers to the QCs satisfying the constraints specified by the load file. In practice, the flexibility in sequencing containers for the YCs exists in the following two aspects: (1) When a container from a certain group is needed for loading, it must be decided on the particular container to retrieve by the YC in the storage yard, with the objective to minimize the rehandle operation; (2) Containers of different groups can be loaded alternately. For example, containers in a yard bay are loaded onto more than one ship bays. Also, there are two loading strategies for the QC, i.e., row wise or column wise. With different QC loading strategies, more flexibility exists in constructing the YC loading sequence.

Scheduling the loading operations for the YCs is to find a better loading sequence given the storage locations of the containers. The objective is to minimize the total completion time of the loading operations. At each epoch, a loading sequence must be constructed for all the containers that will arrive in the following epochs, taking into account the storage locations of those containers. A tabu search algorithm is proposed to solve this problem.

There exist precedence constraints in the loading operations specified by the load file. Some container operations in the storage yard must precede or follow some other operations. Therefore, the feasible solution of a load scheduling is coded into a sequence of containers which are compatible with the precedence constraints. A feasible solution can be defined by a permutation vector $\pi = (\pi(1), \ldots, \pi(n^e))$, where $\pi(x)$ denotes the x^{th} container in π, $n^e = |N^e|$.

4.2.1 Neighborhood structure

Given a feasible solution, π, of the load scheduling, its neighborhood solutions are obtained by swapping two containers (compatible with the precedence constraints) in π. Let $sw(x, x')$ define the operation of swapping the container in position x with the container in position x', $x, x' = 1, 2, \ldots, n^e$. Note that $x \neq x'$. If we denote the new solution obtained from π by applying the swapping move $sw(x, x')$ as π_{sw}, the neighborhood of π, $\boldsymbol{N}(\pi)$, consists of solution π_{sw} by moves sw from a given set OP.

$$OP(\pi) = \bigcup_{\substack{x,x'=1 \\ x \neq x'}}^{n^e} sw(x, x')$$

Thus, we get

$$N(\pi) = \{\pi_{sw} \mid sw \in OP(\pi)\}$$

4.2.2 Tabu list and search strategy

A tabu list is a short time memory of the last several movements. It is applied to have a trace of the evolution of the search to prevent cycling. In our case, once a swap between position x and x' is applied, both position x and x' are added to the tabu list for some time. A swap that involves returning the containers to these positions for a certain number of iterations is not allowed. Let $E = (E_1, E_2, \ldots, E_{maxt})$ be the tabu list of a fixed length $maxt$. $E_t = (\alpha, \beta)$ is a pair element, where α denotes one position, β denotes the other position. E_t, $1 \leq t \leq maxt$, is added to the tabu list in the following way: each element in the tabu list is shifted to the left by one position and E_t is put in position $maxt$, i.e.

$$E_{t-1} = E_t,\ t = 2, \ldots, \max t,$$

$$E_{\max t} = E_t$$

An aspiration criterion is used to override the tabu status of a move. In our algorithm, we allow a tabu move if such a move leads to a neighborhood solution that has a smaller makespan (C_{max}) than that of the best solution obtained so far.

We adopt the best-fit strategy in our search. Given a loading sequence π, its neighborhood is searched using the following strategy: each neighborhood solution of π that is not currently forbidden in the tabu list is evaluated, and then the algorithm moves to the best non-tabu neighbor. That is, the makespan of each solution is calculated, and the solution with the smallest makespan is selected.

4.2.3 Algorithm

The algorithm starts with an initial solution π and a primary empty tabu list E. At each iteration, the neighborhood search strategy is applied and the set of swap operations $OP(\pi)$ is determined. The best non-tabu move sw' is selected, which determines the neighborhood $\pi' = \pi_{sw'}$. Then, the new tabu list E is created. π_{ch}' is set to be the starting solution for the next iteration. The best C_{max}* found and the associated solution π^* are updated. Counters *Iter* and *NIter* trace the number of total iterations and the number of iterations performed without improvement, respectively. The algorithm terminates if there is no improvement over the best solution obtained after a certain number of iterations (*NonImpIter*), or if the total number of iterations reaches a predetermined value *MaxIter*.

The algorithm can be summarized as follows:

Step 1. Initialization: $\pi^* = \pi$, $C_{max}^* = C_{max}(\pi^*)$, $E = \varnothing$, $i = Iter = NIter = 0$.

Step 2. Set $Iter := Iter + 1$, $NIter := NIter + 1$. Find the swap operation $sw' \in$ OP(π), the neighborhood $\pi' = \pi_{sw'}$, and the modified tabu list E. Set $\pi := \pi'$.

Step 3. If $C_{max}(\pi) < C_{max}^*$, then set $\pi^* := \pi$, $C_{max}^* = C_{max}(\pi)$, $Niter := 0$, and go to step 2.

Step 4. If ($Iter \leq MaxIter$) & ($NIter \leq NonImpIter$) then go to step 2; otherwise STOP.

5. EXPERIMENTS

5.1 Input data

To evaluate the solution quality and efficiency of the proposed algorithms, some experiments are conducted and analyzed. The data used in the experiments is generated based on the physical model of a container terminal in the port of Shanghai. In the loading operations, a QC generally loads tens to hundreds containers for a ship. These containers are stacked in blocks in the storage yard. The containers that are destined for the same destination port are often stacked in the same yard bay. Test problems are generated as follows. All times in the experiments are measured in seconds.

- The number of outbound containers in the loading operation varies from 10 to 500; the number of blocks for storing these containers varies from 2 to 5.
- The processing time of a container by an YC is the sum of the operation time required to pick up a container and the transferring time to load the container on an YV. The processing times are generated from a uniform distribution of U(10, 60).
- A rehandle time is needed when the required container is not directly accessible, and is calculated as 15 times the number of containers on the top of the required container.
- For each problem, the containers are randomly classified into different groups. The number of containers in a group is sampled from a uniform distribution of (5, 30) (numbers are rounded). The work schedule of a QC is generated by joining up to 10 container groups in a random sequence.
- A random loading sequence is used to initialize the iterative algorithm.

We use Visual C++ programming language to code the algorithms, and perform the computational test on a personal computer with a Centrino Duo CPU @ 2.0 GHz and 2.0 GB RAM.

5.2 Performance analysis

To evaluate the performance of the iterative search algorithm, the sensitivity analysis is applied to different data settings (number of containers, storage levels of container blocks) and algorithm parameters (number of iterations).

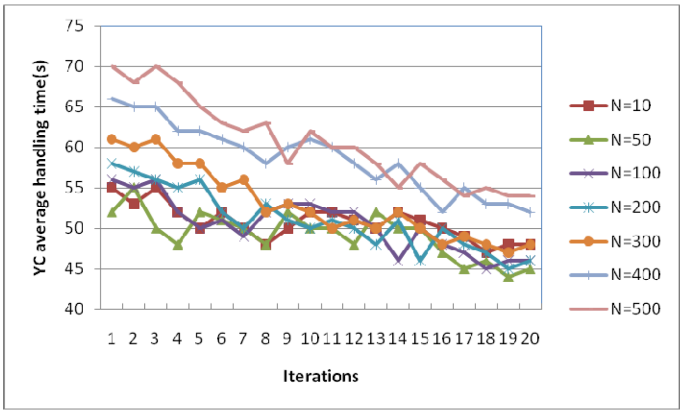


Figure 3. Algorithm performance on YC average handling time

Figure 3 shows that the performance of the algorithm improves with the increase of iterations. 7 sets of problems with different number of containers were generated. Each set has 10 replications. In each replication, a random arrival order of all the containers is generated. And the iterative algorithm is applied to make a decision upon the arrival of each container. It is observed that the reduction of YC average handling time of a container increases when the total number of containers increases (the maximum storage height in the yard is 4 tiers). When the total number of container is 10, the improvement of average handling time is about 12.7% with the use of the iterative search algorithm. The improvement increases to 22.9% when the total number of container is 500. Also the improvement is very limited when the iteration of the algorithm is more than 18.

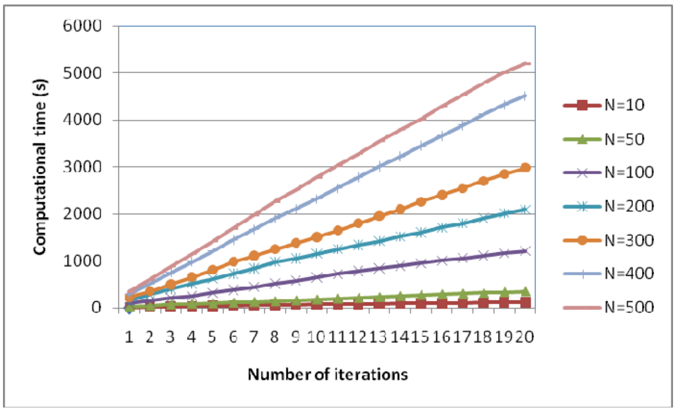


Figure 4. Computational time of the algorithm in the first epoch

Figure 4 illustrates the computational time of the algorithm in the first epoch as a function of the problem size, i.e., e=1, N^e=N. Although the computational time is observed quite long as the increase of the number of containers and the algorithm iterations, this time decreases dramatically for the following epochs (as shown in Figure 5). This phenomenon can be explained by the fact that as the increase of e, good initial loading sequence and storage allocation plan are obtained from the previous epoch.

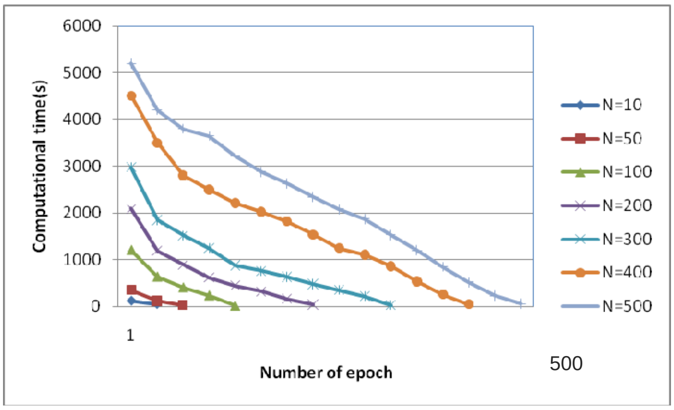


Figure 5. Computational time of the algorithm for the following epochs

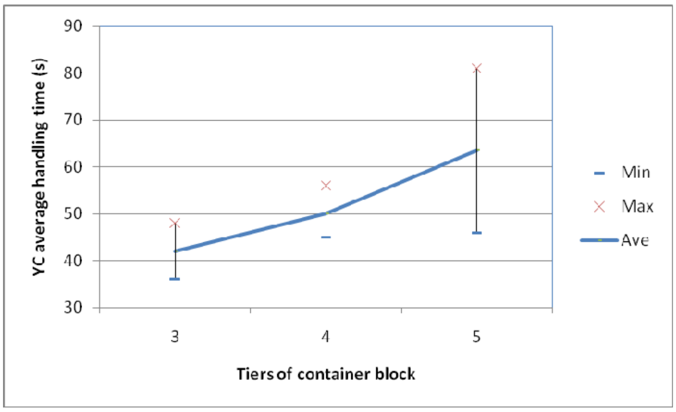


Figure 6. Algorithm performance with different maximum storage height

Sensitivity analysis was performed to analyze the influence of the maximum storage height on the algorithm performance. The results are given in Figure 6. It is shown that when the maximum storage height of container block increases from 3 tiers to 5 tiers, the average handling time in loading operations increases from 42 seconds to 64 seconds (the total number of containers is 100).

6. CONCLUSIONS

This paper proposed a novel iterative search algorithm to solve an integrated problem composed of two sub-problems, namely storage allocation and load scheduling for outbound containers. A GRASP based procedure and a Tabu Search algorithm are developed firstly to solve the two sub-problems independently. Then, the two algorithms are incorporated into a whole in an iterative way: the output of one algorithm is used as the input of the other, capturing the interactions between the two problems. Due to the randomness of the arrivals of outbound containers, the iterative search algorithm is applied to find a better storage location upon the arrival of each container taking into account the loading sequence of all the arriving containers in the future.

The experiments show that the algorithm could find better overall solutions for the two sub-problems. The average YC handling time gets improved as the algorithms continues. However, the improvement becomes quite limited when the iteration of the algorithm gets bigger. And the computational time is reasonable for large sized problems. It was also found that reducing the maximum storage height in the yard resulted in a reduction in the YC handling time.

ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (Project Number: 70802040) and the Natural Sciences and Engineering Research Council of Canada.

7. REFERENCES

Dekker, R., Voogd, P., van Asperen, E. (2006). Advanced methods for container stacking. *OR Spectrum*, 28:563–586.

Glover, F. Future paths for integer programming and links to artificial intelligence. *Computers and Operations Research 1986; 13(5)*; 533-549.

Glover, F. (1989). Tabu search, part I. *ORSA Journal of Computing, 1*: 190–206.

Glover, F. (1990). Tabu search, part II. *ORSA Journal of Computing, 2*: 4–32.

Jung, S.H. and Kim, K.H. (2006). Load scheduling for multiple quay cranes in port container terminals. *Journal of Intelligent Manufacturing, 17*: 479-492.

Kim, K.H. and Kim, K.Y. (1999a). Routing straddle carriers for the loading operation of containers using a beam search algorithm. *Computer & Industrial Engineering, 36*: 109-136.

Kim, K.H. and Kim, K.Y. (1999b). A routing algorithm for a single straddle carrier to load export containers onto a containership. *International Journal of Production Economics, 59*: 425-433.

Kim, K.H. and Kim, K.Y. (1999c). An optimal routing algorithm for a transfer crane in port container terminals. *Transportation Science, 33*: 17-33.

Kim, K.H. and Kim, H.B. (2002). The optimal sizing of the storage space and handling facilities for import containers. *Transportation Research Part B, 36*: 821-835.

Kim, K.H. and Park, K.T. (2003). A note on a dynamic space allocation method for outbound containers. *European Journal of Operational Research, 148(1):* 92-101.

Kim, K.H., Park, Y.M., Ryu, K.R. (2000). Deriving decision rules to locate export containers in container yards. *European Journal of Operational Research, 124(1):* 89-101.

Kozan, E. and Preston, P. (1999). Genetic algorithms to schedule container transfers at multimodal terminals. *International Transaction in Operational Research, 6*: 311-329.

Kozan, E. and Preston, P. (2006). Mathematical modeling of container transfers and storage locations at seaport terminals. *OR Spectrum, 28*; 519-537.

Lee, D.H., Cao, Z., and Meng, Q. (2007). Scheduling of two-transtainer systems for loading outbound containers in port container terminals with simulated annealing algorithm. *International Journal of Production Economics, 107*: 115-124.

Ng, W.C. Crane scheduling in container yards with inter-crane interference. European Journal of Operational Research 2005; 164; 64-78.

Ng, W.C. and Mak, K.L. (2005). Yard crane scheduling in port container terminals. Applied Mathematical Modelling, 29: 263 – 276.

Preston, P., Kozan, E. (2001). An approach to determine storage locations of containers at seaport terminals. Computers & Operations Research, 28: 983-995.

Vis, I.F.A. and De Koster, R. (2003). Transshipment of containers at a container terminal: An overview. European Journal of Operational Research, 147: 1–16.

Stahlbock, R. and Voß, S. (2008). Operations research at container terminals: a literature update. OR Spectrum, 30(1): 1-52.

Steenken, D., Voß S. and Stahlbock, R. (2004). Container terminal operation and operations research --- a classification and literature review. OR Spectrum, 26: 3-49.

Zhang, C., Liu, J., Wan, Y.-W., Murty, K. G, and Linn, R. J. (2003). Storage space allocation in container terminals. Transportation Research Part B, 37: 883-903.

OPTIMIZING BAY DIMENSION IN MIXED PYRAMID STACKING STORAGE SYSTEMS

Dong-Won Jang[1] and Kap Hwan Kim[2]

[1]Department of Industrial Engineering
Pusan National University
Geumjeong-gu 609-735
Busan, Korea
dwjang@pusan.ac.kr

[2] Corresponding author
Department of Industrial Engineering
Pusan National University
Geumjeong-gu 609-735
Busan, Korea
kapkim@pusan.ac.kr

Abstract: Pyramid stacking is a type of block stacking method for cylindrical unit loads such as drums, coils, paper rolls, and so on. This study addresses how to determine the dimension (height and width) of a bay. It is assumed that multiple types of unit loads, which have different retrieval rates and duration of stays from each other, are stored in the same bay. A cost model is proposed to determine the optimal dimension of a bay. A numerical example is provided to illustrate the procedure for the optimization in this paper.

1. INTRODUCTION

Pyramid stacking is a storage method in which cylindrical unit loads are stacked on the floor as shown in Figure 1. Pyramid stacking is one of the storage methods with a high re-handling cost and space utilization. The bay of pyramid stack in Figure 1 consists of 3 tiers by 4 rows at the bottom resulting in 9 unit loads in total. There are 4, 3 and 2 unit loads at each tier, respectively, from the bottom. When a retrieval order is issued for a unit load at a low tier, one or more than one relocation must be performed before the target unit load is retrieved. Such relocations are a major source of inefficiency in handling activities in pyramid stacking systems.

Figure 2 shows the total number of handling each unit load for retrieval in the pyramid stacking system of Figure 1. The k is the index of the tier from the top and l is the index of the position in each tier from the left hand side. The s represents the total number of handlings for retrieving a target unit load from each corresponding position. In case of a unit load at (*3,2*), it requires 4 relocations of unit loads at (*1,1*), (*1,2*), (*2,1*) and (*2,2*) and thus the total number of handlings becomes 5.

For a given number of unit loads in a bay, when the number of unit loads at the lowest tier decreases, the average height of pyramid stacking system must increase, which results in an increase in the expected number of relocations per retrieval. However, the height of a pyramid stacking bay cannot exceed the number of unit loads at the lowest tier because the number of unit loads per tier decreases one by one as the tier goes up. When the number of unit loads at the lowest tier increases, the space required for the bay increases. Park and Kim (2010) attempted to estimate the number of re-handles for a given number of unit loads at the lowest tier and the number of tiers in a bay when all the unit loads are heterogeneous which means that all the unit loads in the bay are different from each other and a retrieval order is issued for a specific unit load in the bay. However, this study extends Park and Kim (2010) to the case where multiple unit loads in the bay are the same type and thus a retrieval order is issued for the unit loads in the same type. Figure 3-(a) illustrates the case where all the unit loads in a bay are different types, while Figure 3-(b) illustrates the case where there are three unit loads in each of three types.

This study derives the expression for the expected number of relocations and suggests a model for determining the optimal dimension of a bay considering the relocation cost and the space cost when there are multiple unit loads of the same type (group) in a bay. Section 2 addresses the case where groups with the same retrieval probabilities. Section 3 analyzes the case where groups have different retrieval probabilities and the retrieval probability is the same as the proportion of the number of unit loads in the bay among different groups. Section 4 treats the most general case where the retrieval probability of a group is not the same as the proportion of the number of the unit loads of the group in the bay. The last section will provides some conclusions.

Figure 1. An example of pyramid stacking in warehouses.

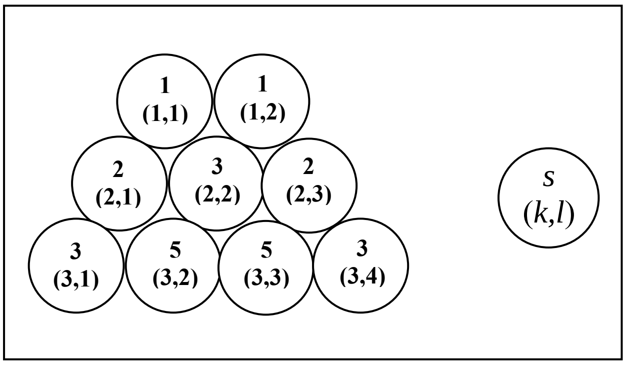


Figure 2. The total number of handlings per unit load for the bay in Figure 1.

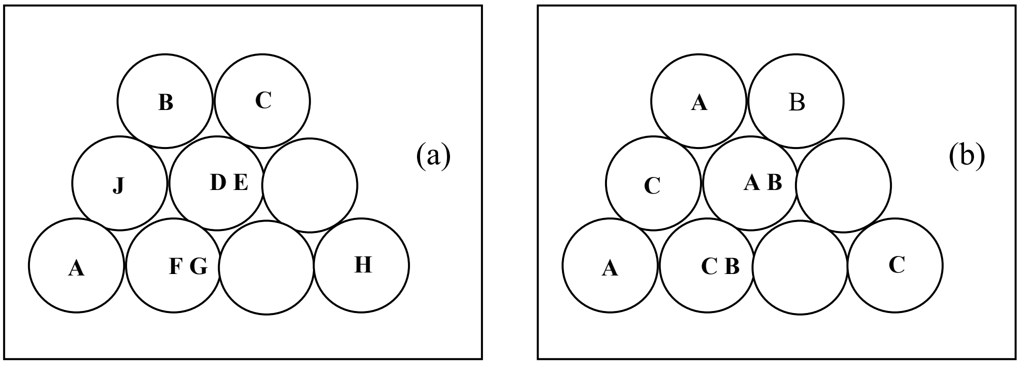


Figure 3. Heterogeneous (a) and homogeneous types of unit loads (b)

Many researchers have analyzed the re-handling operation. Watanabe (1991) analyzed the handling activities in container yards and propose d a simple index, termed as the "index of accessibility" to express the accessibility of a stack considering the number of relocations required to lift a container. Castilho and Daganzo (1993) analyzed the handling activities in inbound container yards. Based on a simulation study, they proposed a formula for estimating the number of relocations for the random retrieval of a container. Kim (1997) proposed a formula for estimating the number of relocations for a random retrieval of an inbound container from a bay. Kim and Kim (1999) analyzed the handling activities for relocations in inbound container yards and used the result for determining the number of devices and the amount of space required for storing inbound containers and for allocating spaces for arriving inbound containers. Kim *et al.* (2000) addressed the problem of locating the arriving outbound containers by considering the weights of the containers. They incorporated the fact that heavier containers are usually stacked in the lowest tiers of bays in vessels; therefore, these must be stacked in the upper tiers at the yard in order to avoid relocations during the loading operations. Kang *et al.* (2006) extended Kim et al. (2000) by relaxing the latter's assumption that the weight information of every arriving container is known with certainty. Kim and Hong (2002) suggested two methods for determining the locations of relocated blocks. Lee and Hsu (2007) and Lee and Chao (2009) discussed how to construct a plan for repositioning export containers within a yard so that no extra re-handles were needed during the loading operation. Park and Kim (2010) compared various stacking methods including block stacking in terms of the expected number of relocations and the storage space. However little is known about the mixed pyramid stacking storage system. As mentioned above, Park and Kim (2010) compared the number of relocations required for picking up a unit load form a stack for various stacking methods. Especially, the analytic results are explained by using examples of cylindrically shaped unit loads, termed as "pyramid stacking" in which unit loads are stacked and centered over the two unit loads in the immediate lower tier and unit loads are lifted from above by overhead bridge cranes. Thieman and Murphy (2000) introduce an automated handling system for coils.

2. THE CASE OF GROUPS WITH THE SAME RETRIEVAL PROBABILITIES

This section derives expressions for the expected number of handlings to retrieve a unit load from a pyramid stacking bay and proposes a method to determine the optimal number of unit loads at the lowest tier in a bay considering the handling cost and the cost of space. The following assumptions are introduced:

① Unit loads of multiple groups are mixed in the same bay.

② A retrieval order is issued for a group of unit loads, which means that any unit load can be retrieved among multiple unit loads of the group to satisfy the retrieval order.

③ If there are multiple unit loads of a group and a retrieval order is issued for the group, then the unit load with the smallest number of handlings will be retrieved. For example, in Figure 2, *s* represents the number of handlings required to pick up the target unit load. This assumption implies that when there is a unit load with *s*=1 among unit loads of the requested group, that unit load will be picked up first. If there is no unit load with *s*=1 of the group, then unit loads with *s*=2 will be picked up. That is, *s* plays the role of the priority for picking up. Thus, we will call *s* "priority" from now on.

④ The number of unit loads in a group stores in a bay is the same for all the groups, that is to say the probability of retrieval of a unit load of a group is the same for all the groups. The last assumption will be modified by each case and will be explained in each section.

The following notations are used for the formulation.

Notation

R	=	*Number of unit loads of the lowest tier in a bay (a decision variable).*
h	=	*Number of tiers required to store all the unit loads. T is the available highest tier considering R in a bay,* $1 \le h \le T, 1 \le T \le R$.
n	=	*Number of slots required to store all the unit loads,* $= \sum_{t=1}^{h} (R-t+1)$.
m	=	*Number of groups of unit loads.*
k	=	*Index representing the tier where the target unit load exist from the top in the bay,* $1 \le k \le h$.
l	=	*Index representing the* l^{th} *unit load at the* k^{th} *tier from left side in the bay,* $1 \le l \le R-h+k$.
$h(s)$	=	*The total number of unit loads for which the number of handling required to pick up the corresponding unit load is s which we call "priority" in this paper.*
$A(k,l)$	=	*The total number of handling for retrieving the* l^{th} *unit load at the* k^{th} *tier in the bay.*
a	=	$max[A(k,l)]$ *for all k and l.*
c_h	=	*Handling cost per unit load.*
r	=	*Arrival rate of retrieval orders per unit time.*
c_s	=	*Cost of ground space for a stack per unit time.*
p	=	*Probability that a specific group of unit loads will be retrieved next, which is the same for all the groups, that is,* $p=1/m$.
$p(s)$	=	*Probability that a group of unit loads, for which the priority for picking up is s, is retrieved from a bay,* $1 \le s \le a$.
$E(s)$	=	*Expected total number of handlings for retrieving a unit load with the priority of s in the bay, which can be expressed as* $s \times p(s)$.
E	=	*Expected total number of handlings for retrieving a unit load in the bay, which can be expressed as* $\sum_{s=1}^{a} E(s)$.
TC	=	*Total cost per unit time.*

The number of handlings for the l^{th} unit load from the left at the k^{th} tier from the top in the stack with the number of unit loads in the lowest tier, *R*, and *h* tier becomes (Park and Kim, 2010)

$$
\begin{aligned}
A(k,l) &= \frac{1}{2}k(1+k) && when \quad k \le l \le R+1-h \\
&= \frac{1}{2}(-R-R^2+2l+2kl+2Rl-2l^2+h+2Rh-2lh-h^2) && when \quad R+1-h \le l \le k \\
&= \frac{1}{2}(1+2k-l)l && when \quad l < k \ \ and \ \ l \le R+1-h \\
&= \frac{1}{2}(1+k+R-l-h)(k-R+l+h) && when \quad l > R+1-t \ \ and \ \ l \ge k \ .
\end{aligned}
$$

When a retrieval order is issued at a pyramid stacking system, it can be assumed that at least one unit load of the corresponding group exists in the bay. The probability that at least one unit load of the requested group exists in a bay is $1-(1-p)^n$. Thus, the probability that a group of unit loads with the priority of *s* is retrieved from the bay can be expressed as follows (Kim, 2010).

$$
\begin{aligned}
p(s) &= \frac{(1-p)^{\sum_{i=1}^{s-1} h(i)} \times \{1-(1-p)^{h(s)}\}}{1-(1-p)^n} && when \quad s \ne 1 \\
&= \frac{1-(1-p)^{h(s)}}{1-(1-p)^n} && when \quad s = 1 .
\end{aligned}
\qquad (1)
$$

Using *p*(*s*) and *s*, we can derive the expected total number of handlings for retrieving a unit load with the priority of *s* from a bay, *E*(*s*). Thus, the expected total number of handlings to pick up a unit load from a bay, *E*, is calculated as follows.

$$
\begin{aligned}
E &= \sum_{s=1}^{a} E(s) = \sum_{s=1}^{a} (s \times p(s)) = \sum_{s=1}^{a} (s \times \frac{(1-p)^{\sum_{i=1}^{s-1} h(i)} \times \{1-(1-p)^{h(s)}\}}{1-(1-p)^n}) && when \quad s \ne 1 \\
&= \sum_{s=1}^{a} E(s) = \sum_{s=1}^{a} (s \times p(s)) = \sum_{s=1}^{a} (s \times \frac{\{1-(1-p)^{h(s)}\}}{1-(1-p)^n}) && when \quad s = 1
\end{aligned}
\qquad (2)
$$

The value of *R* is integral but the value of *h* may not be integral, because there may be cases where unit loads cannot fill all the slots the highest tier in full to store a given number of unit loads. In the case, an interpolation is used for obtaining the value of *E* by using the values of *E* with integral values nearest to *h*.

Using *E* and *R*, we can derive an expression the total cost per unit time. The total cost per unit time, *TC*, can be expressed by the sum of the handling cost per unit time plus the cost of ground space per unit time. The handling cost per unit time becomes rc_hE, while the cost of ground space per unit time is c_sR. Thus, the total cost per unit time becomes:

$$TC = rc_h E + c_s R \qquad (3)$$

By dividing both sides by c_s, a new objective function is obtained.

$$TC = (rc_h / c_s)E + R \qquad (4)$$

There are five groups of unit loads (*m=5*) in case of groups with the same retrieval probabilities. Thus *p=0.2*. Thirty unit loads are stored at a bay (*n=30*). Table 1 shows the average height of stacks for a given number of unit loads, the expected number of handlings per unit load, the handling cost per unit time, and the cost of space per unit time for various numbers of stacks at the lowest tier in the bay.

In this example, the c_s, *r*, and c_h were estimated to be 310.95, 15.02, and 1742.93, respectively. The total cost is found to be minimized when the number of unit loads at the lowest tier is 17 . The highest tier, 2^{nd} tier, in this stacking system is not full up with given the value of *n*.

Table 1. Performance measures for various numbers of stacks at the lowest tier in a bay ($p=0.2$, $n=30$)

R	***h***	***E***	***(rc_h/c_s)E***	***R***	***TC***
8	5	1.867	157.154	8	165.154
9	4	1.507	126.907	9	135.907
10	4	1.341	112.884	10	122.884
11	3	1.239	104.275	11	115.275
12	3	1.170	98.494	12	110.494
13	3	1.124	94.660	13	107.660
14	3	1.093	92.024	14	106.024
15	3	1.071	90.153	15	105.153
16	2	1.051	88.442	16	104.442
***17**	**2**	**1.033**	**86.970**	**17**	**103.970**
18	2	1.021	85.972	18	103.972
19	2	1.013	85.303	19	104.303
20	2	1.008	84.860	20	104.860
21	2	1.005	84.572	21	105.572
22	2	1.002	84.388	22	106.388
23	2	1.001	84.276	23	107.276
24	2	1.000	84.211	24	108.211
25	2	1.000	84.177	25	109.177
26	2	1.000	84.163	26	110.163
27	2	1.000	84.162	27	111.162
28	2	1.000	84.168	28	112.168
29	2	1.000	84.178	29	113.178
30	1	1.000	84.190	30	114.190

3. THE CASE WHERE THE RETRIEVAL PROBABILITIES VARY AMONG GROUOPS

This section addresses the case where the retrieval probabilities vary among groups, derives expressions for the expected number of handlings to retrieve a unit load from a bay, and proposes a method to determine the optimal number of unit loads of the lowest tier in a bay considering the handling cost and the cost of space in the same manner as before. For example, the total number of unit loads in a bay is 24 and the number of unit loads of groups A, B, and C is 12, 8, and 4, respectively. Then the retrieval probabilities of groups A, B, and C become 1/2, 1/3, and 1/6, respectively. Suppose that the probability that a unit load in group g is requested to be retrieved next is represented by p_g. By using p_g, $p(s)$ can be expressed as follows:

$$p(s) = \sum_{g=1}^{m}\left(p_g \times \frac{(1-p_g)^{\sum_{i=1}^{s-1} h(i)} \times \{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^n}\right) \quad when \quad s \neq 1$$

$$= \sum_{g=1}^{m}\left(p_g \times \frac{\{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^n}\right) \quad when \quad s = 1 \qquad (5)$$

$$E(s) = s \times p(s) = s \times \sum_{g=1}^{m}\left(p_g \times \frac{(1-p_g)^{\sum_{i=1}^{s-1} h(i)} \times \{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^n}\right) \quad when \quad s \neq 1$$

$$= s \times p(s) = s \times \sum_{g=1}^{m}\left(p_g \times \frac{\{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^n}\right) \quad when \quad s = 1 \qquad (6)$$

$$E = \sum_{s=1}^{a} E(s) = \sum_{s=1}^{a}(s \times p(s)) = \sum_{s=1}^{a}\left(s \times \sum_{g=1}^{m}\left(p_g \times \frac{(1-p_g)^{\sum_{i=1}^{s-1} h(i)} \times \{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^n}\right)\right) \quad when \quad s \neq 1$$

$$= \sum_{s=1}^{a} E(s) = \sum_{s=1}^{a}(s \times p(s)) = \sum_{s=1}^{a}\left(s \times \sum_{g=1}^{m}\left(p_g \times \frac{\{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^n}\right)\right) \quad when \quad s = 1 \qquad (7)$$

Suppose that there are two groups of unit loads ($m=2$) in a bay, and $p_1=0.7$ and $p_2=0.3$. Thirty unit loads are stored at a bay. The number of required slots is assumed to be 30 ($n=30$). The parameters for TC as have been explained above. Table 2 shows the average height of stacks in a bay, the expected number of handlings per unit load, the handling cost per unit time, and the cost of space per unit time for various numbers of stacks at the lowest tier in pyramid stacking system. It was found the total cost is minimized when the number of unit loads at the lowest in a bay is 11.

Table 2. Performance measures for various numbers of stacks at the lowest tier in a bay ($p_1=0.7$, $p_2=0.3$, $n=30$)

R	***h***	***E***	***$(rc_h/c_s)E$***	***$(c_s / c_s)R$***	***TC***
8	5	1.12	94.61	8	102.61
9	4	1.06	88.88	9	97.88
10	4	1.03	86.85	10	96.85
***11**	**3**	**1.02**	**85.74**	**11**	**96.74**
12	3	1.01	85.16	12	97.16
13	3	1.01	84.81	13	97.81
14	3	1.00	84.59	14	98.59
15	3	1.00	84.45	15	99.45
16	2	1.00	84.35	16	100.35
17	2	1.00	84.28	17	101.28
18	2	1.00	84.24	18	102.24
19	2	1.00	84.22	19	103.22
20	2	1.00	84.20	20	104.20
21	2	1.00	84.20	21	105.20
22	2	1.00	84.19	22	106.19
23	2	1.00	84.19	23	107.19
24	2	1.00	84.19	24	108.19
25	2	1.00	84.19	25	109.19
26	2	1.00	84.19	26	110.19
27	2	1.00	84.19	27	111.19
28	2	1.00	84.19	28	112.19
29	2	1.00	84.19	29	113.19
30	1	1.00	84.19	30	114.19

4. THE CASE WHERE THE RETRIEVAL PROBABILITY IS NOT THE SAME AS THE STORAGE PROBABILITY

This section addresses the case where the retrieval probability is not the same as the storage probability, which is the most general case. The following additional notation is used for the formulation.

d_g = *Average duration of stay of unit loads in group g.*

f_g = *Retrieval frequency per unit time for group g.*

p_g = *Probability that a unit load in a bay is included in group g.*

q_g = *Probability that a retrieval request is for group g.*

Using the data on the average dwell time and retrieval frequency for group g, p_g can be estimated as follow.

$$p_g = d_g \times \frac{f_g}{\sum_{g=1}^{m}(d_g \times f_g)} \quad (8)$$

The retrieval probability of group g can also be expressed as:

$$q_g = \frac{f_g}{\sum_{g=1}^{m} f_g} \quad (9)$$

Using the estimated probabilities for storage and retrieval, we can be expressed as follows.

$$p(s) = \sum_{g=1}^{m}\left(q_g \times \frac{(1-p_g)^{\sum_{i=1}^{s-1} h(i)} \times \{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^n}\right) \qquad when \quad s \neq 1 \quad (10)$$

$$= \sum_{g=1}^{m}\left(q_g \times \frac{\{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^{n}}\right) \qquad when \quad s = 1$$

$$E(s) = s\times p(s) = s\times\sum_{g=1}^{m}\left(q_g \times \frac{(1-p_g)^{\sum_{i=1}^{s-1} h(i)} \times\{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^{n}}\right) \qquad when \quad s \neq 1$$

$$= s\times p(s) = s\times\sum_{g=1}^{m}\left(q_g \times \frac{\{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^{n}}\right) \qquad when \quad s = 1 \qquad (11)$$

$$E = \sum_{s=1}^{a} E(s) = \sum_{s=1}^{a}(s\times p(s)) = \sum_{s=1}^{a}\left(s\times\sum_{g=1}^{m}\left(q_g \times \frac{(1-p_g)^{\sum_{i=1}^{s-1} h(i)} \times\{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^{n}}\right)\right) \qquad when \quad s \neq 1$$

$$= \sum_{s=1}^{a} E(s) = \sum_{s=1}^{a}(s\times p(s)) = \sum_{s=1}^{a}\left(s\times\sum_{g=1}^{m}\left(q_g \times \frac{\{1-(1-p_g)^{h(s)}\}}{1-(1-p_g)^{n}}\right)\right) \qquad when \quad s = 1 \qquad (12)$$

Table 3 shows the data for the duration of stay and retrieval frequency for groups of unit loads. Using the expressions (8) and (9), the storage and retrieval probabilities for each group can be estimated. In this example, $p_1=0.29$, $p_2=0.46$, and $p_3=0.26$. Also, $q_1=0.22$, $q_2=0.44$, and $q_3=0.33$.

Table 3. Data for the duration of stay and retrieval frequency for groups.

Group	***Duration of stay***	***Retrieval frequency***
A 5		2
B 4		4
C 3		3

There are three groups of unit loads ($m=3$). The number of required slots is assumed to be 30($n=30$). The parameters for *TC* have mentioned before. Table 4 shows the average height of stacks in a bay, the expected number of handlings per unit load, the handling cost per unit time, and the cost of space per unit time for various numbers of stacks. It was found the total cost is minimized when the number of unit loads at the lowest tier is 13.

Table 4. Performance measures for various numbers of stacks at the lowest tier in a bay

R	***h***	***E***	***$(rc_h/c_s)E$***	***$(c_s / c_s)R$***	***TC***
8 5		1.33	111.95	8	119.95
9 4		1.16	97.37	9	106.37
10 4		1.09	91.92	10	101.92
11 3		1.06	88.91	11	99.91
12 3		1.04	87.26	12	99.26
***13 3**		**1.02**	**86.24**	**13**	**99.24**
14 3		1.02	85.58	14	99.58
15 3		1.01	85.15	15	100.15
16 2		1.01	84.81	16	100.81
17 2		1.00	84.56	17	101.56
18 2		1.00	84.40	18	102.40
19 2		1.00	84.31	19	103.31
20 2		1.00	84.25	20	104.25
21 2		1.00	84.22	21	105.22
22 2		1.00	84.20	22	106.20
23 2		1.00	84.19	23	107.19
24 2		1.00	84.19	24	108.19
25 2		1.00	84.19	25	109.19
26 2		1.00	84.19	26	110.19
27 2		1.00	84.19	27	111.19
28 2		1.00	84.19	28	112.19
29 2		1.00	84.19	29	113.19
30 1		1.00	84.19	30	114.19

5. CONCLUSIONS

This study discussed a relocation problem in which unit loads in a pyramid stacking system are classified into multiple groups and any unit load in a group may be picked up when a retrieval order is for the group. This study addressed how to estimate the number of relocations during the retrieval process in various situations. Three cases were analyzed: the case where the retrieval probability is the same for all items; the case where the retrieval probabilities of items differ across multiple groups but are proportional to the number of unit loads in each group stored in a bay; the case where the retrieval probabilities of different groups are different from each other and are not proportional to the number of stored unit loads of each group in the pyramid stacking system. Numerical examples were performed to illustrate the optimizing method.

ACKNOWLEDGEMENTS

This work was supported by the Grant of the Korean Ministry of Education, Science, and Technology" (The Regional Core Research Program/Institute of Logistics Information Technology)

6. REFERENCES

Castilho, B., Daganzo, C.F. (1993). Handling strategies for import containers at maritime terminals. *Transportation Research Part B, Vol. 27, No. 2, 151-166.*

Kang J., Ryu, K. R. and Kim, K. H. (2006). Determination of storage locations for incoming containers of uncertain weights. *Ali, A. and Dapoigny (Eds.): IEA/AIE 2006, LNAI 4031, 1159-1168.*

Kim, K. H. (1997). Evaluation of the number of rehandles in container yards. *Computers & Industrial Engineering, Vol. 32, No. 4, 701-711.*

Kim, K. H., Hong, G. P. (2006). A heuristic rule for relocating blocks. *Computers & Operations Research, Vol. 33, No. 4, 940-954.*

Kim, K.H., Kim, H.B. (1999). Segregating space allocation models for container inventories in port container terminals. *International Journal of Production Economics, Vol.59, No.1-3, 415-423.*

Kim, K.H., Park, Y.M. and Ryu, K.R. (2000). Deriving decision rules to locate export containers in container yards. *European Journal of Operation Research, Vol.124, No.1, 89-101.*

Kim, S.W (2010). *A statistical analysis on the number of re-handles in mixed storage yards.* Dept. of Logistics Information Technology Graduate School, PNU, Korea.

Lee, Y., Chao, S.-L. (2009). A neighborhood search heuristic for pre-marshaling export containers. *European Journal of Operational Research, Vol. 196, 468-475.*

Lee, Y., Hsu, N.-Y.(2007). An optimization model for the container pre-marshaling problem. *Computers & Operations Research, Vol. 34, 3295-3313.*

Park, T.-K., Kim, K.H. (2010). Comparing handling and space costs for various types of stacking methods. *Computers & Industrial Engineering, 58:501-508.*

Thieman, M. J., Murphy, R. L. (2000). Automated coil handling and in-process inventory system. *Steel Technology(November), 29-31.*

Watanabe, I.R.(1991). Characteristics and analysis method of efficiencies of container terminal an approach to the optimal loading/unloading method. *Container Age, March, 36-47.*

Session C3: Seaport and Transportation 3

·Day1: Sep. 15 (Wed.)

·Time: 15:00 - 16:20

·Chair: Hans-Dietrich Haasis

·Room: Iris, 4F

LOGMS

The 1st International Conference on Logistics and Maritime Systems

September 15th–17th, 2010 Busan, Korea

A MULTI-PRODUCT DYNAMIC INBOUND ORDERING AND SHIPMENT SCHEDULING PROBLEM AT A THIRD-PARTY WAREHOUSE

Byung Soo Kim[1] ,and Woon-Seek, Lee[2*]

1. Department of Industrial & Systems Engineering
Aub urn University
Auburn, AL, 36849-5346, U.S.A.

2. Department of Systems Management & Engineering
Pukyong National University
Busan 608-739, Korea
Corresponding Author : Email : iewslee@ pknu.ac.kr

Abstract: This paper considers a dynamic inbound ordering and shipment scheduling problem for multiple products that are transported from a supplier to a warehouse by common freight containers. The following assumptions are made: (i) each ordering in a period is immediately shipped in the same period, (ii) the total freight cost is proportional to the number of containers used, and (iii) demand is dynamic and backlogging is not allowed. The objective of this study is to identify effective algorithms that simultaneously determine inbound ordering lot-sizes and a shipment schedule that minimize the total cost consisting of ordering cost, inventory holding cost, and freight cost. This problem can be shown in NP-hard, and this paper presents a heuristic algorithm that exploits the properties of an optimal solution. Also, a shortest path reformulation model is proposed to obtain a good lower bound. Simulation experiments are presented to evaluate the performance of proposed procedures.

1. INTRODUCTION

For the couple of decades, the reduction of transportation cost and warehousing cost have been two important issues to enhance logistic efficiency and demand visibility in a supply chain. The logistic alliances and specialized Third-Party-Logistic (TPL) providers have been growing to reduce the costs in industry. In a dynamic planning time period, the issue of transportation scheduling for inbound ordering and shipping of products to TPL warehouse by proper transportation modes at right time and the issue of lot size dispatching control including inventory control to the customers have become significantly important for production and distribution management. Each warehouse purchases multiple products and uses a freight container as a transportation unit to ship its purchased (or manufactured) products to retailers, which may lead to the managerial decision problems including lot-sizes for each product, container types used, loading policy in containers, and the number of containers used. Thus, this provides us with a motivation to investigate the optimal lot-sizing and shipment scheduling problem. Also, the managerial decision problems have arisen in TPL.

Several articles have attempted to extend the classical Dynamic Lot-Sizing Model (DLSM) incorporating production-inventory and transportation functions together. Hwang and Sohn (1985) investigated how to simultaneously determine the transportation mode and order size for a deteriorating product without considering capacity restrictions on the transportation modes. Lee (1989) considered a DLSM allowing multiple set-up costs consisting of a fixed charge cost and a freight cost, in which a fixed single container type with limited carrying capacity is considered and the freight cost is proportional to the number of containers used. Fumero and Vercellis (1999) proposed an integrated optimization model for production and distribution planning considering such operational decisions as capacity management, inventory allocation, and vehicle routing. The solution of the integrated optimization model was obtained using the Lagrangean relaxation technique. Lee *et al.* (2003) extended the works of Lee (1989) by considering multiple heterogeneous vehicle types to immediately transport the finished product in the same period it is produced. It is also assumed that each vehicle has a type-dependent carrying capacity and the unit freight cost for each vehicle type is dependent on the carrying capacity. Lee *et al.* (2003) considered a dynamic model for inventory lot-sizing and outbound shipment scheduling in the Third-Party Warehousing domain. They presented a polynomial time algorithm for computing the optimal solution. Jaruphongsa et al. (2005) analyzed a dynamic lot-sizing model in which replenishment orders may be delivered by multiple shipment modes with different lead times and cost functions. They proposed a polynomial time algorithm based on the dynamic programming approach. However, the aforementioned works have not considered a multiple product problem.

Anily and Tzur (2005) considered a dynamic model of shipping multiple items by capacitated vehicles. They presented an algorithm based on a dynamic programming approach. Norden and Velde (2005) dealt with a multiple product problem of determining transportation lot-sizes in which the transportation cost function has piece-wise linear as to a transportation capacity reservation contract. They proposed a Lagrangean relaxation algorithm to compute lower and upper bounds. This paper analyzes a dynamic inbound ordering and shipment scheduling problem for multiple products that are transported from a supplier to TPL warehouse by common freight containers in a supply chain. It is assumed that each order is not allowed to split to ship in the different periods and shipped immediately in the same period and the total freight cost is proportional to the number of containers used. Further, no backlogging is allowed. We extend to the multi-product problem in this paper. The main objective of this study is to simultaneously determine the lot-sizes and the shipment schedule that minimize the total cost which consists of ordering, inventory holding, and freight costs.

A mixed integer optimization model is developed in section 2 followed by the characterization of the structure of an optimal solution in section 3. A shortest path formulation of the problem is developed to obtain a good lower bound in section 4. A two-phased heuristic algorithm is then presented in section 5 to exploit the properties of an optimal solution. Computational results from a set of simulation experiments are also presented to evaluate the performance of the proposed procedures in section 6. Finally, concluding remarks follow in the last section.

2. MODEL FORMULATION

The following notations are defined to formulate the problem:

- T = length of the time horizon,
- t = *time index* ($t = 1,2,\cdots,T$),
- M = number of products,
- i = product index ($i = 1,2,\cdots,M$),
- d_{ti} = amount demanded for product i in period t,
- $M_{ti} = \sum_{k=t}^{T} d_{ki}$,
- W = carrying capacity of a container,
- S_i = setup cost of product i in period t ,
- h_i = unit inventory holding cost of product i from period t to period $t+1$,
- F = unit freight cost of container in period t ,
- x_{ti} = amount of product i ordered and shipped by container in period t ,
- y_t = number of containers used in period t (nonnegative integer),
- I_{ti} = amount of inventory of product i at the end of period t , and
- z_{ti} = 1, if a setup is incurred for product i in period t , and 0, otherwise.

We assume that all products have the same weight and volume specifications. The objective of the problem is to determine (x_{ti}, y_t) for $t = 1,2,\cdots,T$ and $i = 1,2,\cdots,M$ so that all demands are satisfied at the minimum total cost. Therefore, the *T*-period problem can be formulated as a mathematical program as follows:

(*P*)
$$\underset{x_{ti},y_t}{Minimize} \sum_{t=1}^{T} \{ \sum_{i=1}^{M} S_i \cdot z_{ti} + \sum_{i=1}^{M} h_i \cdot I_{ti} + F \cdot y_t \} \quad (1)$$

$$s.t. \quad I_{ti} = I_{t-1,i} + x_{ti} - d_{ti}, \qquad \forall t,\ i, \quad (2)$$

$$\sum_{i=1}^{M} x_{ti} \le W \cdot y_t, \qquad \forall t, \quad (3)$$

$$x_{ti} \le M_{ti} \cdot z_{ti}, \qquad \forall t, i, \quad (4)$$

$$I_{0i} = I_{Ti} = 0, \qquad \forall i, \quad (5)$$

$$x_{ti} \ge 0,\ I_{ti} \ge 0, \qquad \forall t, i, \quad (6)$$

$$y_t : \text{nonnegative integer}, \qquad \forall t. \quad (7)$$

The constraint (3) implies that the total ordering amount is restricted by the total carrying capacity associated with the number of containers used in the period. The constraints (2) through (7) define a closed bounded convex set and the objective function is concave, so that the problem attains its minimum at an extreme point of the convex set. The extreme points will be further characterized in association with the optimal solution in the next section.

3. OPTIMAL SOLUTION PROPERTIES

The model (*P*) may be represented by a network model as shown in Figure 1. In the network, two flow types are defined as follows:

(1) The aggregate flow is defined as the flow between node 0 and nodes $(1,2,\cdots,T)$.

(2) The individual flow is defined as the flow between nodes $(1,2,\cdots,T)$ and nodes $((1,1), (1,2),\cdots,(T,M))$.

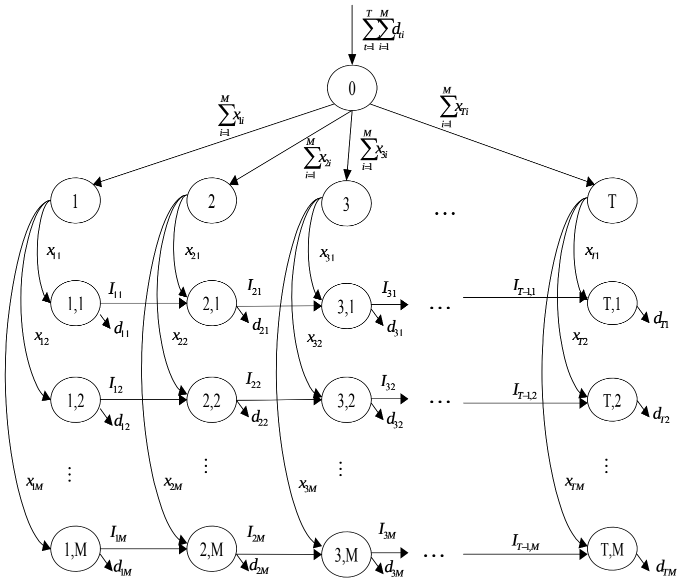


Figure 1. Network representation of the model *P*

Here, the arcs in the aggregate flow are restricted by the capacities associated with the number of containers used whereas the arcs in the individual flow are not. An optimal solution of the model (*P*) occurs at extreme points. In a network theory, such an extreme point can be interpreted as an extreme flow (refer to Florian et al. (1971)). In a network without arc capacities, a feasible flow is an extreme flow if it does not have a positive loop. Also, in a network with arc capacities, a feasible flow is an extreme flow if and only if each loop has at least one saturated arc. In Figure 1, loops can be formed by the following two cases:

(1) Between the aggregate flow and the individual flow, for example, the loop can be formed by the sequences of nodes (0), (1), (1,1), (2,1), (2), and (0).

(2) On the individual flows, for example, the loop can be formed by the sequences of nodes (1,1), (1), (1,*M*), (2,M), (2), (2,1), and (1,1).

The ordering point, partial shipment point, and inventory point are defined as follows to further examine the optimal solution properties of the problem:

(1) Period *t* is a ordering point for product *i* if $x_{ti} > 0$.

(2) Period *t* is a partial shipment point if $nW < \sum_{i=1}^{M} x_{ti} < (n+1)W$ (*n* is a nonegative integer).

(3) Period *t* is an inventory point for product *i* if $I_{ti} = 0$.

The optimal solution property of Wagner and Whitin (1958), i.e., $x_{ti} \cdot I_{t-1,i} = 0$, is always the extreme flow for the model (*P*). Such solutions prevent the formation of positive loops. If $x_{ti} \cdot I_{t-1,i} \neq 0$, on the other hand, the properties of theorems 1 and 2 must be satisfied to have an extreme flow.

Theorem 1. In the model (P), the optimal solution has at most one partial shipment point between two consecutive inventory points for product i.
Proof. Suppose that there exists an optimal solution that has two partial shipment points between two consecutive inventory points for product i. If this is the case, a loop may be formed by the sequence of nodes (0), (1), (1,1), (2,1), (2), and (0) in Figure 1. At least one of the arcs (0,1) and (0,2) must be saturated for this feasible flow to be the extreme flow. This feasible flow is not an extreme flow since these arcs are not saturated. Therefore, the proof is completed.

Theorem 2. In the model (P), the optimal flow should not form the positive loop in the individual flow.
Proof. Suppose that there exists a feasible flow that satisfies the property of Theorem 1 and has a loop formed by the sequence of nodes (1,1), (1), (1,2), (2,2), (2), (2,3), (3,3), (3), (3,1), (2,1), and (1,1) in Figure 1. Since the arcs in the individual flow are not capacitated, there exists a positive loop formed by unsaturated arcs. This feasible flow is not an extreme flow. Therefore, the proof is completed.

Unfortunately, the problem (P) is NP-hard, which is mainly due to the existence of unbounded integer variables, $\{y_t\}$, even though the variable is restricted by total demands divided by the unit carrying capacity of container (W). To tackle this problem, a shortest path reformulation model is proposed to a good lower bound in the next section.

4. SHORTEST-PATH REFORMULATION MODEL

It is obvious that the problem (P) may potentially be a very large Mixed Integer Linear Programming (MILP) problem. Therefore, an efficient formulation is required to make optimization viable for large-scale problems. In this respect, the network formulation approach proposed by Eppen and Martin (1987) may be adapted to analyze the problem (P).

Our formulation utilizes a network representation of single product lot-sizing problems. The lot-sizing problem for each product is represented by the graph $\boldsymbol{G}$=($\boldsymbol{N}$, $\boldsymbol{A}$). The node set is $\boldsymbol{N}$={t} where the index t indicates the time period. The arc set is $\boldsymbol{A}$={(s, t) : $s<t$} where (s, t) corresponds to ordering of item i in period s with the lot size $\sum_{k=s}^{t-1} d_{ki}$ which satisfies the demand in periods $s,\cdots,t-1$. Figure 2 depicts a graphical representation of the graph, $\boldsymbol{G}$, for a particular item. Define $d_{ist}=\sum_{k=s}^{t} d_{ki}$ ($t \ge s$) and $c_{itk}=\sum_{s=t}^{k-1} h_i \cdot d_{i,s+1,k}$.

Let w_{itk} represent the flow along the arc (t, k) in the graph for item i. Also let the parameter b_i be $\min\{t : d_{ti} > 0\}$. Then, the problem (P) may be formulated as the shortest path problem (SP) on the graph $\boldsymbol{G}$ as follows:

$$(SP) \qquad Minimize \sum_{i=1}^{M}\sum_{t=1}^{T}\{\sum_{k=t}^{T} c_{itk}\cdot w_{itk} + S_i \cdot z_{it}\} + \sum_{t=1}^{T} F \cdot y_t \quad (8)$$

$$s.t. \quad \sum_{i=1}^{M}\sum_{k=t}^{T} d_{itk}\cdot w_{itk} \le W \cdot y_t, \quad \forall t, \quad (9)$$

$$\sum_{t=1}^{b_i}\sum_{k=t}^{T} w_{itk} = 1, \quad \forall i, \quad (10)$$

$$\sum_{k=1}^{t-1} w_{i,k,t-1} - \sum_{k=t}^{T} w_{itk} = 0, \quad \forall i, t > b_i, \quad (11)$$

$$\sum_{k=t}^{T} w_{itk} \le z_{it}, \quad \forall i,t, \quad (12)$$

$$w_{itk} \ge 0, \quad \forall i,t,k \ge t, \quad (13)$$

$$z_{it} \in \{0,1\}, \quad \forall i,t, \text{ and} \quad (14)$$

$$y_t : \text{nonnegative integer}, \quad \forall t. \quad (15)$$

Constraint (9) implies that total ordering amount is restricted by total carrying capacity associated with the number of containers used in the period. Constraints (10) and (11) represent the flow balance constraints. Constraint (10) enforces the boundary condition while constraint (12) maintains the flow balance. Constraint (13) establishes the integrality of setups even when the flow variables are fractional. But, there exist unbounded integer variables, $\{y_t\}$, which obstruct a significant increase in solution speed for the problem (SP). To resolve this, an LP relaxation model of the problem (*SP*) is proposed to efficiently obtain a good lower bound to the problems (*P*) and (*SP*).

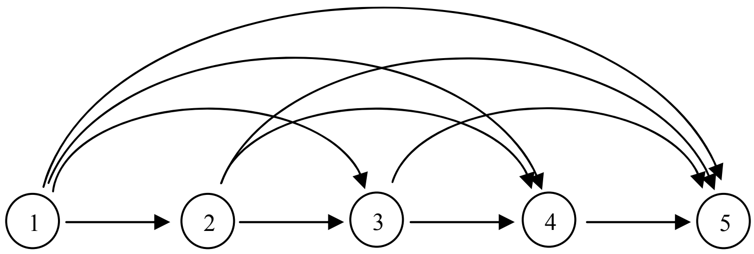


Figure 2. Graph ***G*** for the problem *P* with 5-periods

As suggested by Eppen and Martin (1987), the problem (*SP*) may be modified by adding slack variables u_{it} and v_t, and adjusting objective function coefficients. The fixed cost S_i and the marginal shipment cost *F/W* are used as the objective function coefficients corresponding to u_{it} and v_t, respectively. Also, a coefficient of 0 is used for y_t while the coefficient of c^1_{itk} is used for w_{itk} in the objective function, where $c^1_{itk} = S_i + F/W \cdot d_{itk} + \sum_{s=t}^{k-1} h_i \cdot d_{i,s+1,k}$.

The cost coefficient (17) includes setup, marginal shipment, and inventory holding costs that are incurred to order and ship the cumulative demands of item *i* from period *t* to period *k*. Consequently, an LP relaxation model to the problem (*SP*) can be written as follows:

$$(SPR) \qquad Minimize \sum_{i=1}^{M}\sum_{t=1}^{T}\{\sum_{k=t}^{T} c^1_{itk} \cdot w_{itk} + S_i \cdot u_{it}\} + \sum_{t=1}^{T}(F/W)\cdot v_t \quad (16)$$

$$s.t. \qquad \sum_{i=1}^{M}\sum_{k=t}^{T} d_{itk} \cdot w_{itk} = W \cdot y_t, \quad \forall t, \quad (17)$$

$$\sum_{k=t}^{T} w_{itk} + u_{it} = z_{it}, \quad \forall i,t, \quad (18)$$

$$\sum_{s=1}^{t} y_s \ge L_t, \quad \forall t, \quad (19)$$

$$\sum_{s=t}^{T} y_s \le U_t, \quad \forall t, \quad (20)$$

$$u_{it} \ge 0, \quad \forall i,t, \quad (21)$$

$$v_t \ge 0, \quad \forall t, \quad (22)$$

$$w_{itk} \ge 0, \quad \forall i,t,k \ge t, \quad (23)$$

$$0 \le z_{it} \le 1, \quad \forall i,t, \quad (24)$$

$$y_t \ge 0, \quad \forall t, \quad (25)$$

and (10) and (11).

where $\langle x \rangle$ = smallest integer greater than or equal to x, $L_t = \left\langle \sum_{i=1}^{M}\sum_{s=1}^{t} d_{si}/W \right\rangle$ and $U_t = \left\langle \sum_{i=1}^{M}\sum_{s=t}^{T} d_{si}/W \right\rangle + (T-t+1)$.

Constraints (19) and (20) enforce the boundary condition leading to a significant increase in solution speed for the problem (SPR). A heuristic algorithm to solve the model (P) presented above is developed based on the properties of Theorems 1 and 2 and presented in the next section.

5. TWO-PHASED HEURISTIC ALGORITHM

A heuristic algorithm to efficiently find a good solution consists of two phases. In the first phase, a feasible solution is generated by forwardly performing a marginal cost coefficient approach. In the second phase, an adjustment mechanism is backwardly performed to obtain a good solution closer to the optimal solution.

5.1 Phase 1 : marginal cost coefficient approach

A marginal cost coefficient $M_i(t)$ is defined by

$$M_i(t) = \frac{S_i + H_i(t-1) - h_i \cdot (t-1)^2 \cdot d_{ti}}{t(t-1)}, \quad \text{where} \quad H_i(t) = h_i \sum_{k=1}^{t} (k-1) d_{ki}, \tag{26}$$

as the work of Dixon and Silver (1981). If $M_i(t)$ is positive, cost reductions are expected by including d_{ti} in the present lot. Otherwise, cost reductions are not expected. A feasible solution generated by this approach satisfies the optimal solution property of Wagner and Whitin (1958), which does not consider the carrying capacity of containers used. Thus, feasible solutions based on the marginal cost coefficient approach may cause an excess use of containers that are not fully but partially loaded. It may also lead to an increase in total cost. Based upon the property of Theorem 1, the total transportation amount is adjusted to fully utilize the carrying capacity of containers employed. A marginal cost coefficient approach to obtain a feasible solution is introduced in Appendix.

5.2 Phase 2 : Adjustment mechanism

The marginal cost coefficient approach may easily offer a feasible solution, but the solution may be a weak solution in comparison with the optimal solution. An adjustment mechanism needs to be performed to obtain a better solution than the feasible solution generated by the marginal cost coefficient approach. The adjustment mechanism consists of three steps. In the first step, the ordering amount of an item in a specific ordering point is merged into the immediate previous ordering point if a cost reduction is expected, which is backwardly performed for all items. In the second step, the transportation amount in a specific partial shipment point is transferred into the immediate previous partial shipment point if a cost reduction is expected, which is backwardly performed for all periods. However, an excess inventory may be accumulated in the first period because the first and second steps are only backwardly performed over the planning horizon. Thus, a part of the excess inventory amount in the first period can be transferred into the next partial shipment point in the final third step. While the adjustment mechanism is performed, an additional use of the container, an increase in ordering points, and inventory shortages are not allowed. The details are introduced in Appendix.

6. COMPUTATIONAL RESULTS

To analyze the performance of the proposed procedures, simulation experiments have been conducted with the following data:

(1) Set M = 3, 6, 8, 10 and T = 4, 6, 8, 12, 18.

(2) Demands are generated from a normal distribution $N(\mu_i, \sigma_i^2)$, where μ_i is generated from a uniform distribution U(25, 100) and σ_i is either μ_i or $\mu_i / 5$ with equal likelihood.

(3) A setup cost is given by $S_i = TS_i^2 \cdot \mu_i / 2$ and TS_i =1, 3, 6, where TS_i denotes EOQ time supply.

(4) Set $h_i = 1$ and W = 100, 200, 300 and their respective freight costs are $F = i \cdot W$, for i = 1, 2, 3.

A Visual Basic 6.0 computer code is developed and run on an IBM PC to perform the proposed procedures. In addition, CPLEX 6.0.2 package is run to find the optimal solution. However, an optimal solution has not been identified for large-size test problems with greater than 8 periods and 6 items, which is mainly due to excessive number of nodes and memory consumption generated by the branching mechanism of CPLEX package, and the limitations on computing capability. For most small-size test problems, however, an optimal solution has been found. A tight lower bound to the problem (P) has also been found by solving the problem (SPR). Solutions to the problem (SPR) are then compared with the optimal and heuristic solutions by computing the following measures:

(1) The gap between the optimal and heuristic solutions, *gap*1: $\dfrac{Z_{Heu} - Z_{Opt}}{Z_{Opt}} \times 100$,

(2) The gap between the optimal and *SPR* solutions, *gap*2: $\dfrac{Z_{Opt} - Z_{LB}}{Z_{Opt}} \times 100$,

(3) The gap between the heuristic and *SPR* solutions, *gap*3: $\dfrac{Z_{Heu} - Z_{LB}}{Z_{LB}} \times 100$, and

where Z_{Opt} = objective value of the optimal solution, Z_{Heu} = objective value of the heuristic solution, and Z_{LB} = objective value of the SPR solution.

Four observations have been collected for each combination of input parameters, and the average gaps are calculated and summarized in Tables 1, and 2. Table 1 compares the performance of heuristic with the optimal solution (*gap1*) and the performance of *SPR* solution with the optimal solution (*gap2*) for small-size test problems. For the *gap1* in the table, the heuristic differs on the average by 5.00% from the optimal solution for small-size test problems with less than or equal to 8 periods and 6 items. It should also be noted that the performance of the heuristic tends to get worse as the carrying capacity of container increases. For the *gap*2 in the table, we observed *SPR* offers a good solution within 1.33% in comparison with the optimal solution for small-size test problems. It shows that *SPR* offers a tight lower bound to the problem (*P*).

Table 2 shows the performance of the heuristic algorithm in comparison with the *SPR* solution for large-size test problems. It can be seen that the heuristic differs on the average by 7.52% from the *SPR* solution for large-size test problems. It is expected, from these results, that the heuristic may still be efficient for more large-size problems.

Finally, the average computational times (in seconds) taken by the heuristic is shown that the heuristic offers the solution within the maximum of 0.11 seconds. Meanwhile, the average computational times (in seconds) taken by the CPLEX package significantly increase as *T* and *M* increase. It is expected that the heuristic become much more efficient as *T* and *M* increase.

Table 1. Average gap between the heuristic and best solutions and gap between the optimal solution and SPR solution

M	*W*	*F*	*gap*1(%)			*gap2*(%)		
			T=4	*T*=6	*T*=8	*T*=4	*T*=6	*T*=8
3	100	100	1.3559	2.7874	5.1777	1.1772	1.6417	1.6702
		300	1.7898	3.0841	4.7167	0.8176	1.0987	1.4691
		600	2.8119	3.8635	5.4159	0.5527	0.7346	0.9847
	200	200	2.0729	2.8319	4.2319	2.5957	2.5099	2.7379
		600	3.0850	5.9507	8.0789	1.7726	1.7763	2.2171
		1200	8.3334	6.3752	8.8390	1.2019	1.1854	1.4950
	300	300	3.6065	7.6909	7.7756	2.3785	3.4845	3.0520
		900	4.8704	6.3391	8.8308	1.5825	2.3286	2.0203
		1800	7.0486	14.0629	9.7020	1.0541	1.5555	1.3407
6	100	100	0.6196	1.1704	1.4564	0.4611	0.6523	0.7427
		300	1.2941	2.5847	2.6645	0.3117	0.4342	0.5068
		600	2.2322	2.6575	3.7306	0.2099	0.2892	0.3392
	200	200	1.9965	2.9142	3.2117	1.2222	1.4244	1.3180
		600	4.3898	5.6813	6.2075	0.9052	0.9500	0.9981
		1200	5.4869	6.3088	7.0597	0.6131	0.6336	0.6708
	300	300	4.1554	3.4063	4.1705	1.6003	1.9790	2.4937
		900	6.4958	5.3064	7.2493	1.0724	1.3131	1.7372
		1800	9.9681	8.4883	8.8710	0.7183	0.8730	1.1623
Average			5.0094			1.3346		

Table 2. Average gap between heuristic and *SPR* solutions

M	*W*	*F*	*gap*3(%)		*M*	*W*	*F*	*gap*3(%)	
			T=12	*T*=18				*T*=12	*T*=18
3	100	100	9.7367	10.0765	6	100	100	3.1382	5.418
		300	8.6651	9.0246			300	4.0628	5.8183
		600	9.1313	9.5808			600	5.4116	6.2762
	200	200	8.5241	11.1934		200	200	5.0804	5.0204
		600	11.7213	13.2555			600	7.3723	5.9033
		1200	10.8434	11.9022			1200	7.2719	6.2702
	300	300	13.6019	13.586		300	300	6.6017	8.2617
		900	15.7267	14.3807			900	7.6924	9.1432
		1800	11.5861	12.9157			1800	10.2084	11.6122
8	100	100	3.6983	5.2464	10	100	100	3.0139	4.0546
		300	3.8619	4.7230			300	2.9610	3.6677
		600	3.5615	4.1164			600	2.9759	3.4995
	200	200	5.7693	6.5905		200	200	4.0106	4.7290
		600	6.4921	6.5333			600	5.6894	5.4801
		1200	6.5983	5.5789			1200	5.9424	5.5155
	300	300	8.4312	7.3722		300	300	5.4264	6.7891
		900	9.2417	8.9414			900	8.4028	8.4187
		1800	9.3788	10.2278			1800	9.0933	9.0505
Average									7.5153

7. CONCLUDING REMARKS

This paper investigates a dynamic inbound ordering and shipment scheduling problem for multiple products that are transported from a supplier to TPL warehouse by common freight containers. Since the problem is NP- hard, a two-phased heuristic is proposed based on the optimal solution properties. The problem has also been formulated as a shortest path model to obtain a tight lower bound. To evaluate the performance of proposed procedures, computational results from a set of simulation experiments are also presented. It is shown, on the average, that the heuristic offers a good solution within 5.00% in comparison with the optimal solution for small-size test problems. It is also discussed that the shortest path model offers a tight lower bound with the average gap of 1.33% in comparison with the optimal solution for small-size test problems. Further, the heuristic differs on the average by 7.52% from the solution to the shortest path model for large-size test problems. With regard to computational time, the heuristic is so much faster than CPLEX package for all factors of T, M, and W, and will become much more efficient as T and M increase. We expect that the heuristic may be efficiently employed for large-size problems in the real world.

ACKNOWLEDGEMENTS

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF—2008-313-D01222).

REFERENCES

Anily, S. and Tzur, M., 2005. Shipping multiple items by capacitated vehicles: an optimal dynamic programming approach. *Transportation Science 39(2):* 233-248.

Dixon, P. and Silver, E. A., 1981. A heuristic solution procedure for the multi-item single-level, limited capacity, lot-sizing problem. *Journal of Operations Management 2(1):* 1710-1720.

Eppen, G. D. and Martin, R. K., 1987. Solving multi-item capacitated lot-sizing problems using variable redefinition. *Operations Research 35(6):* 832-848.

Florian, M., Rossin-Arthiat, M., and D., DeWerra, 1971. A property of minimum concave cost flows in capacitated networks. *INFOR 9*: 293-304.
Fumero, F. and Vercellis, C., 1999. Synchronized development of production, inventory, and distribution schedules. *Transportation Science 33(3):* 330-340.
Jaruphongsa, W., Centinkaya, S., and Lee, C. Y., 2005. A dynamic lot-sizing model with multi-mode replenishments: polynomial algorithms for special cases with dual and multiple modes. *IIE Transactions 37:* 453-467.
Hwang, H. and Sohn, K. I., 1985. An optimal policy for the dynamic transportation-inventory model with deteriorating items. IIE Transactions 17, 233-241.
Lee, C. Y.,1989. A solution to the multiple set-up problem with dynamic demand. *IIE Transactions 21(3):* 266-270.
Lee, C. Y., Centinkaya, C., and Jaruphongsa, W., 2003. A dynamic model for inventory lot sizing and outbound shipment scheduling at a third-party warehouse. *Operations Research 51(5):* 735-747.
Lee, W. S., Sox, C. R., and Kim, C. H., 2003. A dynamic production and transportation model with heterogeneous vehicle types. *International Journal of Industrial Engineering 10(4),* 420-426..
Lippman, S. A., 1969. Optimal inventory Policy with multiple set-up costs. *Management Science 1(1):* 118-138.
Norden, L. and Velde, S., 2005. Multi-product lot-sizing with a transportation capacity reservation contract. *European J. of Operational Research 165:* 127-138.
Wagner, H. M. and Whitin, T. M., 1958. Dynamic version of the economic lot size model. *Management Science 5(1):* 89-96.

APPENDIX

Phase 1 : marginal cost approach

Step 1. Set t=1 and generate $M \times T$ demand matrix, in which the element (i,t) denotes d_{ti}.

Step 2. If t=T, $x_{Ti} = d_{Ti}$ for all i and stop the procedure. Otherwise, $x_{ti} = d_{ti}$ for all i and if all demand in period t are zero, go to step 5.Step 3.

Step 3. For all i and t ($2 \le t \le T$), compute $M_i(t)$.

Step 3.1 For all i, include demands associated with the largest positive $M_i(t)$ in the present lot, i.e., $x_{ti} = x_{ti} + d_{ti}$. If lot sizes for all products are determined, go to step 4. Otherwise, for the next largest positive $M_i(t)$, repeat step 3.1.

Step 4. Adjust the total transportation amount in each period.

Step 4.1 If $\sum_{i=1}^{M} x_{ti} \le W$, go to step 5. Otherwise, compute $MOD(\sum_{i=1}^{M} x_{ti}, W)$, where $MOD(p, q)$ denotes the remainder of p divided by q.

Step 4.2 If $MOD(\sum_{i=1}^{M} x_{ti}, W) \ge W - MOD(\sum_{i=1}^{M} x_{ti}, W)$, go to step 5.

Step 4.3 For demand d_{ti} associated with the smallest positive $M_i(t)$, compute $\varepsilon = \min\{MOD(\sum_{i=1}^{M} x_{ti}, W), d_{ti}\}$.

Step 4.4 If $\varepsilon = MOD(\sum_{i=1}^{M} x_{ti}, W)$, compute $x_{ti} = x_{ti} - MOD(\sum_{i=1}^{M} x_{ti}, W)$ and go to step 5.

Step 4.5 If $\varepsilon = d_{ti}$, compute $x_{ti} = x_{ti} - d_{ti}$ and $\sum_{i=1}^{M} x_{ti} = \sum_{i=1}^{M} x_{ti} - d_{ti}$. Repeat step 4.3 for the next smallest positive $M_i(t)$.

Step5. For each product, subtract demands included in the present lot from the demand matrix. Set t=t+1 and go to step 2.

Phase 2 : adjustment mechanism

To describe the adjustment mechanism, the following notations are introduced:

$\langle x \rangle$ = the minimum integer exceeds x and $(x)^+ = \begin{cases} x, & if \ x \ge 0 \\ 0, & otherwise \end{cases}$

Step 1. $M_0 = \{1, 2, \cdots, M\}$.

Step 1.1. If $M_0 = \{\phi\}$, go to step 2. Otherwise, evaluate the following:

$i_1 =$ product index which has the maximum setup cost in M_0, $PP(i_1) =$ set of ordering points for the product i_1, $l_1 = \max\{t \mid t \in PP(i_1)\}$ and $l_2 = \max\{t \mid t \in PP(i_1) - \{l_1\}\}$

Step 1.2. If $l_2 = 0$, compute $M_0 = M_0 - \{i_1\}$ and go to step 1.1. Otherwise, compute an SR as follows:

$$SR = S_{i_1} - h_{i_1} \cdot (l_1 - l_2) \cdot x_{l_1,i_1} + \left[\left\langle \sum_{j=1}^{M} x_{l_1,j} \Big/ W \right\rangle - \left\langle (\sum_{j=1}^{M} x_{l_1,i_1} - x_{l_1,i_1}) \Big/ W \right\rangle \right] \cdot F$$

$$- \left[\left\langle (\sum_{j=1}^{M} x_{l_2,j} + x_{l_1,i_1}) \Big/ W \right\rangle - \left\langle \sum_{j=1}^{M} x_{l_2,j} \Big/ W \right\rangle \right] \cdot F.$$

Step 1.3. If $SR > 0$, set $x_{l_2,i_1} = x_{l_2,i_1} + x_{l_1,i_1}$, $x_{l_1,i_1} = 0$, $PP(i_1) = PP(i_1) - \{l_1\}$ and go to step 1.1.

Otherwise, set $PP(i_1) = PP(i_1) - \{l_1\}$ and go to step 1.1.

Step 2. $t_1 = T$.

Step 2.1. Compute the following:

$\varepsilon_t = MOD\left(\sum_{i-1}^{M} x_{t,i}, W \right)$ for $1 \le t \le t_1$, $PS = \{t \mid \varepsilon_t > 0, 1 \le t \le t_1\}$, $t_1 = \max\{t \mid t \in PS\}$, and

$t_2 = \max\{t \mid t \in PS - \{t_1\}\}$. If $t_1 = 1$, go to step 3.

Step 2.2. If $\varepsilon_{t_1} + \varepsilon_{t_2} > W$, set $t_1 = t_2$ and go to step 2.1.

Step 2.3. If $\sum_{i=1}^{M} x_{t_1,i} - \varepsilon_{t_1} > 0$, go to step 2.4. Otherwise, compute an FR as follows:

$$FR = F + \sum_{i=1}^{M} S_i \cdot \delta(x_{t_1,i}) - \sum_{i=1}^{M} h_i \cdot (t_1 - t_2) \cdot x_{t_1,i} - \sum_{i=1}^{M} S_i \cdot (\delta(x_{t_1,i}) - \delta(x_{t_2,i}))^+$$

If $FR > 0$, set $x_{t_2,i} = x_{t_2,i} + x_{t_1,i}$, $x_{t_1,i} = 0$, for all i, $PS = PS - \{t_1\}$ and go to step 2.1.

Othe rwise, set $PS = PS - \{t_1\}$ and go to step 2.1.

Step 2.4. Compute the following :

$H_{CUM} = 0$, $\hat{x}_{t,i} = x_{t,i,}$ for $t = t_1, t_2$ and all i,

$M_1 = \{i \mid \hat{x}_{t_1,i} > 0, \ \forall i\}$ and $M_2 = \{i \mid x_{t_1,i} \le \varepsilon_{t_1}, \ i \in M_1\}$.

Step 2.5. if $M_2 = \{\phi\}$, go to step 2.6. Otherwise, compute the following:

$HC(i) = S_i \cdot \delta(\hat{x}_{t_1,i}) - h_i \cdot (t_1 - t_2) \cdot \hat{x}_{t_1,i} - S_i \cdot (\delta(\hat{x}_{t_1,i}) - \delta(\hat{x}_{t_2,i}))^+$, $i \in M_2$,

$HC = \max\{HC(i), i \in M_2\}$ and $i_2 =$ inde x associated with HC,

$\hat{x}_{t_2,i_2} = \hat{x}_{t_2,i_2} + \hat{x}_{t_1,i_2}$, $\varepsilon_{t_1} = \varepsilon_{t_1} - \hat{x}_{t_1,i_2}$, $\hat{x}_{t_1,i_2} = 0$.

If $\varepsilon_{t_1} > 0$, compute the following :

$M_1 = M_1 - \{i_2\}$, $H_{CUM} = H_{CUM} + HC$, $M_2 = \{i \mid \hat{x}_{t_1,i} \le \varepsilon_{t_1}, i \in M_1\}$, and go to step 2.5.

Otherwise, set $PS = PS - \{t_1\}$ and go to step 2.1.

Step 2.6. Compute as following:

$FR(i) = F + H_{CUM} + S_i \cdot \delta(\hat{x}_{t_1,i}) - h_i \cdot (t_1 - t_2) \cdot \hat{x}_{t_1,i} - S_i \cdot (\delta(\hat{x}_{t_1,i}) - \delta(\hat{x}_{t_2,i}))^+$, $i \in M_1$,

$FR = \max\{FR(i), i \in M_1\}$ and $i_3 =$ inde x associated with FR.

If $FR > 0$, go to step 2.7. Otherwise, set $PS = PS - \{t_1\}$ and go to step 2.1

Step 2.7. Compute the following:

$\hat{x}_{t_2,i_3} = \hat{x}_{t_2,i_3} + \varepsilon_{t_1}$, $\hat{x}_{t_1,i_3} = \hat{x}_{t_1,i_3} - \varepsilon_{t_1}$, $x_{t,i} = \hat{x}_{t,i}$, for $t = t_1, t_2$ and all i, and

$PS = PS - \{t_{i_1}\}$. And go to step 2.1.

Step 3. Compute the following:

$\varepsilon_t = MOD(\sum_{i=1}^{M} x_{t,i}, W)$ for $1 \le t \le T$, $PS^* = \{t \mid \varepsilon_t > 0, 1 \le t \le T\}$ an d $t_3 = \min\{t \mid t \in PS^*\}$, and $M_3 = \{i \mid x_{1,i} > 0\} \cap \{i \mid x_{t_3,i} > 0\}$.

If $t_3 = 1$ or $M_3 = \{\phi\}$, stop the adjustment mechanism. Otherwise, compute the following:

$e_{t_3} = MOD(\sum_{i=1}^{M} x_{t_3,i}, W)$, $I_{\min}(i) = \min\{\sum_{t=1}^{s}(x_{t,i} - d_{t,i}),\ 1 \le s \le t_3 - 1,\ W - \varepsilon_{t_3}\},\quad i \in M_3$, and

$H_{\max} = \max\{h_i \cdot I_{\min}(i),\ i \in M_3\}$ and $i^* =$ index associated with $H_{\max}$.

Step 3.1. If $H_{\max} > 0$, compute the following:

$x_{1,i^*} = x_{1,i^*} - I_{\min}(i^*)$, $x_{t_3,i^*} = x_{t_3,i^*} + I_{\min}(i^*)$ and stop the adjustment mechanism. Otherwise, stop the adjustm ent mechanism.

A FRAMEWORK FOR SYSTEMATIC DESIGN AND OPERATION OF CONDITION-BASED MAINTENANCE SYSTEMS: EVIDENCE FROM A CASE STUDY OF FLEET MANAGEMENT AT SEA PORTS

Marco Lewandowski[1*] ,and Bernd Scholz-Reiter[1]

[1]BIBA Bremer Institut für Produktion und Logistik GmbH
at the University of Bremen
Hochschulring 20
Bremen, Germany
e-mail: lew@biba.uni-bremen.de

Abstract: Ongoing improvement of logistics and intermodal transport leads to high requirements regarding availability of machine resources like straddle carriers or gantry cranes. Accordingly, efficient maintenance strategies for port equipment have to be established. The change to condition-based maintenance strategies promises to save resources while enhancing availability and reliability. This paper introduces a framework of methods and tools that enable the systematic design of condition-based maintenance systems on the one hand and offers integrated support for operating such systems on the other hand. The findings are evaluated based on a case-study of a German seaport and illustrate the usage of the system based on managing the equipping process of machines with sensors for condition monitoring as well as bringing the system into the operation phase.

1. INTRODUCTION

Global distributed production structures and according supply networks have to work efficiently. Especially sea ports and transhipment terminals, the backbone of Europe's economy, have to posses lean structures and ensure seamless integration in the supply chain. The ongoing improvement of logistics and intermodal transport with respect to through-put time and cost reductions shall satisfy future demands for scalable structures in times of economic growth and recession. Accordingly, this leads to high requirements regarding efficient maintenance strategies for port equipment like straddle carriers or gantry cranes.

While cyclic and reactive maintenance actions are still the representative method in practice, the change to condition-based monitoring of equipment is ongoing. Condition-based maintenance itself promises to make maintenance processes more efficient through decentralized decision units in terms of cognitive sensor applications at certain crucial components for instance so that the machine itself will be able to trigger a maintenance action in terms of automated control and cooperation. Hence, this paper presents the exploration of a fleet management case in a German seaport in which the specific requirements of the design and operating phase were examined and transferred to an adopted systematic procedure model and a methodology framework.

The paper is organized as follows: The first chapter introduces the maintenance topic with the specific requirements regarding port equipment. A state of the art review refers to topical work on condition-based maintenance in general and according endeavours to build a comprehensive framework for such systems. Chapter two presents the methodologies which are part of a framework that enables and supports the design and operation of condition-based maintenance systems on top of existing assets. Its application based, on a case study at a German seaport, verifies the applicability in chapter three. The last chapter presents a conclusion on the work done and gives an outlook for necessary further research and work to be done to put such systems into practice.

1.1 Maintenance in General

The term maintenance describes the combination of all technical and administrative actions that have to be fulfilled in order to retain the functioning condition of a technical system or to restore it to a state in which it can perform in the required manner. To this end, the main aim of maintenance is to secure the preferably continuous availability of machines. Based on this definition the holistic view on the maintenance topic is clear. The several processes based on the typical maintenance tasks as presented in the following table 1 consequently require task-specific know-how.

Table 1. Typical maintenance tasks

Service	Actions that have to be done in order to retain proper performance, generally by replacing worn out parts.
Inspection	Actions to determine the present state of a system
Repair	Actions that re-initiate proper performance of a technical system after an unscheduled breakdown for instance
Improvement	The combination of technical and administrative actions that enhance the performance of a technical system (without changing its function)

Due to physical reasons or based on cost-benefit-considerations, maintenance-free systems can generally not be guaranteed within the whole life cycle so that planned reserves for wear lead to recovering or replacement of components. Maintenance is generally considered a cost factor whereas the amount depends ultimately on the chosen maintenance strategy. One distinguishes generally between run-to-failure and preventive strategies (Mobley, 2004).

Run-to-failure maintenance, also referred to as corrective maintenance, includes the general strategy that maintenance actions are fulfilled only after the recognition of an error in order to repair the specific system. In the narrower sense this is not a real strategy due to a missing long-term planning perspective. Accordingly, the breakdown of the machine and the disruption of the processes are accepted. In practice this kind of strategy has a comparatively minor interest but is still used regarding low-value assets that do not endanger the overall function of a machine and can be changed or repaired very quickly.

In contrast to the corrective strategy, preventive maintenance should take effect before a failure or dysfunction occurs and should accordingly decrease the probability of such events. Actions can be triggered generally according to two principles, which are based on fixed cyclic time criteria or based on predefined measurement criteria such as the remaining reserve of wear. Cyclic maintenance strategies are most commonly used in practice. That means that reserves of wear and other maintenance activities are planned and dimensioned in the research and development phase so that a maintenance action will be done in a cyclic manner each after a specific period of days or certain operating hours.

An alternative preventive strategy focuses on condition monitoring, which means that the topical condition of a component is the decision criteria for a maintenance action. Focusing on condition-based monitoring in the contemporary understanding, modules and components are elementary equipped with sensors that aim at appraising the condition of these elements by statistically analyzing the sensor data. First implementations and pilots focusing mainly on the algorithms that calculate the remaining life time are present now (Schuh *et al.*). In case of any (predefined) abnormalities, corresponding actions can be fulfilled to rebuild the condition of the particular machine (cp. for instance Byrne *et al.*, 1995; Mobley, 2002; Higgins *et al.*, 2008). The impulse for this process chain is consequently given by the machine itself.

1.2 Specific Requirements for Maintenance at Seaports

First implementations of condition-based maintenance systems are present mainly in stationary production plants, e.g. in the processing industry. The methodology and findings of this paper have been developed focusing on maintenance in seaports, essentially the maintenance of straddle carriers. Besides the general aim to establish a methodology and toolset for the design and operation of condition-based management systems – at this state independent of the specific application – the underlying scenario has additional specific requirements.

In this context, the mobile application on vehicles is a special challenge in the face of technologically feasible and economic solutions. At present, in the logistical environment at seaports, machines and vehicles generally are maintained in cyclic manner. Machines and installations operate in rough environments. Compared to high-technology machinery in production environments, systems in logistics have often less investment costs, so that technology for condition monitoring has to be chosen seriously, especially if systems are supplementarily equipped. Accordingly, the following table gives an overview of the special situation compared to standard application of predictive maintenance in plant operations for instance.

Table 2. Exemplary comparison of predictive maintenance application scenarios

Criteria	**Predictive maintenance in logistics**	**Standard application of predictive maintenance (e.g. plant operation)**
Installation	Mobile	Stationary
Environment	Rough and dirty	Depends on the application but in general comparably cleaner
Value of the asset(s)	Comparably low	Comparably high
Power supply	24 V board supply, batteries, accumulator batteries	In general not restricted
Working time	Up to 24 hours, utilization variable	Up to 24 hours, in general constant
Communication	Wireless	In general not restricted, partly existent
Sensors	In general not existent, systems engineering necessary	Partly existent
Data quality	Noise due to drive, motion or environmental impacts	In general constant

1.3 State of the art

Optimizing the efficiency of maintenance processes is a broad research field regarding several tasks of industrial engineering, e.g. product life cycle management as a holistic approach to gain efficient product life phases in face of costs and feedback loops between phases for further improvement in new product generations.

More detailed predictive maintenance in the sense of regular monitoring of the actual mechanical condition aims at using and analyzing sensor information for a condition-based maintenance program and scheduling. Among others, this includes essentially the following aspects

- Non-destructive measurement of physical parameters
- Assessment of components' condition
- Prediction of next failure time
- Replacement strategies and scheduling

Existing sensors which enable non-destructive measurement include for instance several physical principles. Among others these are vibration monitoring, thermography or sonography (cp. for instance Mobley, 2002). Latest sensor developments focus on online oil analysis in hydraulic cycles or engines (e.g. Macián *et al.*, 2003). However, research and development regarding non-destructive measurement is ongoing and will still offer a crucial basis for further applications of condition monitoring. As an example Jomdecha *et al.* (2009) presents the development of a sensor that is able to measure the condition of steel ropes. The ropes at gantry cranes and straddle carriers at seaports are security relevant and significant in face of total management costs so that available sensors could cut these costs, increase reliability and efficient change strategies.

The assessment of components' condition based on the acquired data goes along with the signal processing of this raw data. State-of-the-art methods are, for instance, time series analysis based on statistical principles such as autoregression and moving averages (ARMA). While these were originally developed to analyze and predict the behaviour of economic markets (e.g. stock markets) a transfer to technical diagnostic has been fulfilled (e.g. Pham *et al.*, 2010). Moreover spectral analysis based on fourier transformations enables the interpretation of vibration characteristics.

Based on condition assessment the prediction of next failure time is the following step for using the basic sensor information in maintenance management. For this, mathematical methods mainly based on times series analysis and predictions are used. Both linear statistical models such as autoregressive moving average (ARMA) or autoregressive integrated moving average (ARIMA) and nonlinear time series analysis based on neural networks are the most common approaches.

Using such information in maintenance management systems requires the integration in replacement and scheduling strategies. Focusing on the operative replacement strategies and the scheduling of resources, the aim is to determine a replacement strategy which is - under consideration of cost and risk aspects – optimal or at least economical. Kans *et al.* (2008) for instance simplifies decision making by proposing a common database structure to store maintenance information and acquired data. However, this is the most unsolved task in relation to condition-based monitoring. This means, that processes that are triggered, based on locally distributed decisions by a machine itself lead to dynamic and complex system behaviour, whereas standard deterministic and global planning and controlling approaches show difficulties (Scholz-Reiter *et al.*, 2009). Decentralized auto-controlled algorithms for instance based on multi-agent technology promise to gain more efficiency in solving these tasks.

Focusing on the engineering process as a systematic design of condition-monitoring systems, engineering models have to be taken into consideration. However, such guidelines have to be adapted to the new scenarios. Some requirements were figured out by Lewin (1995) who focuses on the selection of critical components for a condition monitoring.

2. METHODOLOGY FRAMEWORK

While cyclic and reactive maintenance are still the representative strategy in practice, the change to condition-based monitoring of equipment is ongoing and accordingly requires an aligned systems engineering model to support development and integration of such systems. The superior framework accordingly supports all phases of the condition monitoring system, reaching from the design, the development and the integration over the starting phase and operation to the feedback at the end-of-life – the latter to improve further product generations. The following Figure 1 shows an overview of phases that are required to build a condition monitoring system and operate it as suggested in this paper.

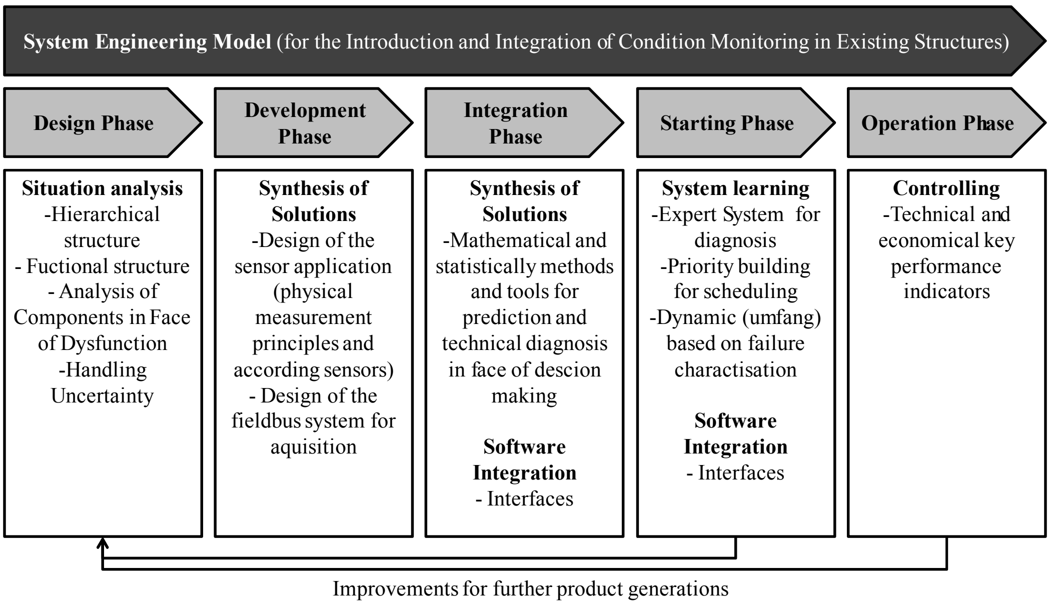


Figure 1. Framework of requirements to build a condition monitoring system based on different phases

Each phase has its specific requirements for the project engineer which furthermore leads to specific methods and tools that have to be applied to design and operate a condition monitoring system. It is not relevant whether an existing or a new product will be equipped. The total methodology framework is based on a user-centred design approach which means that the user-specific requirements of project engineers and the operative maintenance mangers are taken into consideration. The detailed approaches for the different phases are presented below.

2.1 System Engineering Model

A general system engineering approach serves as a basis for analyzing the as-is state, weak points and requirements as well as developing the later to-be state. However, cost-effective executions of such projects have to substantiate the particular procedure of the model.

The main requirements of the model are

- Modelling as-is structure and finding weak points (critical components)
- Finding and configuring the right sensors to monitor condition of components

- Finding and configuring the mathematical methods to analyse condition of components
- Calculating or determining replacement strategies

Among others, these aspects are represented according to the phases of the framework and described in detail in the following parts. More Details can be found in Lewandowski *et al.* (2009).

2.2 Design and Development Phase

On the basis of a user-centred design the methodology framework supports the analysis of the product situation (essentially regarding an existing object). This is based on a software-supported analysis tool. The tool mainly focuses on analyzing the as-is-state of the object based on the hierarchical and functional structure of all components. Moreover, each component can be characterized based on its priority. This is based on dysfunction probability, costs, security aspects, etc. Reports on this topic accordingly enable developing a measurement system for components that have a high priority within the system as a hole.

Within integrated software a database of possible sensors is provided to equip components with specific sensor application. Parameterizations are part of the system which describe the normal vs. the warning operation range of a components' attribute. All technical parameters for the field bus technology which connect the sensors directly at the machine are set within this software application.

2.3 Integration Phase

The integration phase consists of a methodology for configuring the prognosis and technical diagnosis based on the sensor time series. The necessary prerequisite should be the definition of a normal operating range for a sensor-equipped component and accordingly an additional warning and critical range. Thus, saved sensor data are analyzed based on this definition so that critical states can be defined from today's point of view.

Moreover, time series are analyzed based on state-of-the-art algorithms such as ARMA or ARIMA as well as neural networks so that future values can be predicted. Using this in relation with the operating range definition, one is able to predict the state of a component within different time intervals such as the next 24 operating hours.

The technical diagnosis is based on a neural network approach which classifies the possible causes for failures based on the sensor states. This means that the input consist of information per sensor that defines whether a component is in normal, warning or critical state. The neural network as a black box system delivers concrete information such as 'Failure in Oil Cooling System'. The interconnections of the methods are presented in the following Figure 3 which also shows the respective end-user perspective.

2.4 Starting Phase

The starting phase follows the initial operation of the whole system and is necessary for configuring the system based on the technical diagnosis. This means, that neural networks or other expert systems that are used according to the integration phase have to be trained with respect to the later operation phase. Efficient software-based user interfaces should support these tasks. This aims to externalize the users' knowledge and know-how.

Additional tasks of the starting phase focus on configuring the algorithms for the later usage of condition monitoring data in the planning and controlling of operative maintenance tasks. For that purpose a priority building based on the failure messages should be completed and the dynamic amount of maintenance actions have to be calculated and scheduled. In line with methodology framework a genetic algorithm has been developed to schedule maintenance tasks based on information of the condition monitoring system.

2.5 Operation Phase

Main outcome from the previous steps is a condition based maintenance system that operates according to the standards of the Open System Architecture for Condition Based Maintenance (OSA-CBM) (Swearingen *et al.*, 2007) and which consists of layers that include at least sensor module, signal processing, condition monitor, health assessment and prognostics. However, the operating phase of the system also enables the acquisition of key performance indicators based on economical and technical parameters. Here it is necessary that information from the operating phase flows back to the planning tools to enhance the database and accordingly to improve the system. Report possibilities include the statistical analysis of failures according to the respective measurements of the condition monitoring system. New product generations can profit from this enhanced data pool.

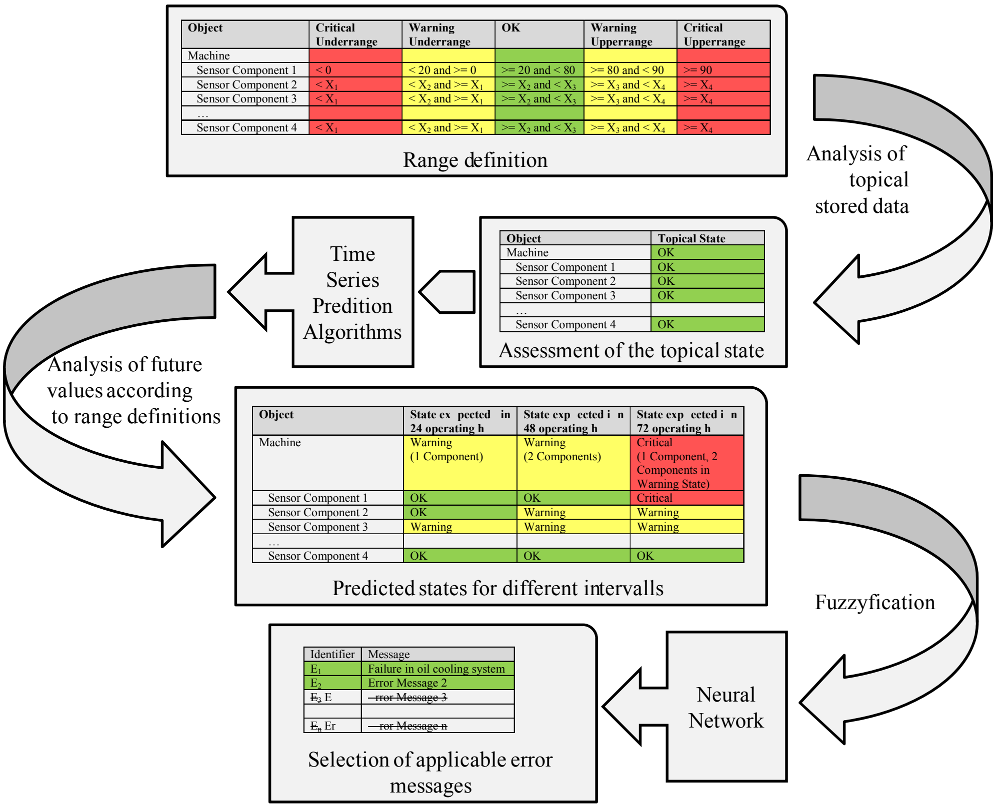


Figure 3. Interconnections between methods and tools that enable condition monitoring in the integration and starting phase

3. FINDINGS

The findings include consequently a framework mainly based on software and hardware for planning purposes, data acquisition, storage, and processing as well as illustration and connection to enterprise systems, e.g. for planning and controlling maintenance resources. In this context the developments have been tested with port equipment at a German seaport. The case study was based on a single straddle carrier on which condition monitoring techniques have been applied retrospectively and independently from the manufacturer.

The hardware system, especially sensors and fieldbus technology, is based mainly on standard industry sensor technology (e.g. temperature, pressure and quality sensors for the hydraulic system, vibration sensors, etc.) and an Ethernet-based fieldbus as known in industry automation. Data acquisition is done by an Onboard-Unit based on an Industrial PC running Windows CE. Data is transferred wireless via GSM or WLAN to a server component which is responsible for storing data, analyzing it and presenting it to the user. An additional software application is responsible for the project management analyzes the system and configures the applications as described in the previous chapter. The software architecture is illustrated in the following Figure 4.

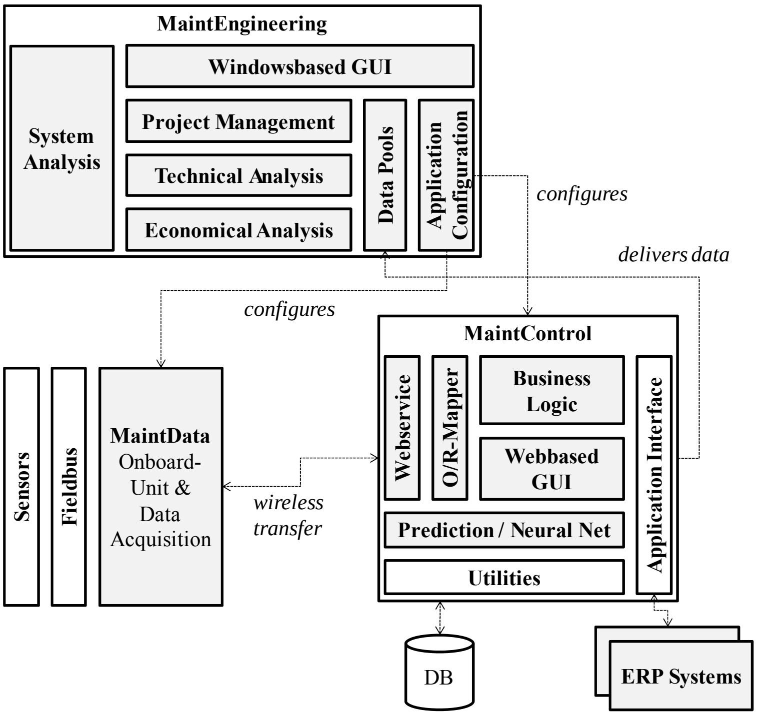


Figure 4. Architecture of the developed solution

From the users point of view so called MaintEngineering is a windows-based software application which supports the project eng ineer in designing th e cond ition m onitoring application by su pporting the system analysis in cluding technical and econom ical aspects. A data pool (e.g. i ncluding a database o f a vailable sens ors) su pports t he sy stem specification and the configuration of other software components.

While MaintData runs as a terminal service on the Onboard-Unit, so called MaintControl is the central component for the end-user. The component includes storage of the data, business logic to link information, prediction algorithms, and application of decision rules to finally present it to the user. This is done based on a webbased user interface.

4. CONCLUSION AND OUTLOOK

In conclusion, the developed approach serves as a practice-oriented example to organize the change to condition-based maintenance strategies. It addresses at least two main problems of maintenance management essentially with a focus on condition-based m onitoring and according strategies. In m ost exi sting sce narios a cha nge to c ondition based maintenance strategies requires an aligned system engineering process that integrates the design and operating phase as proposed here.

At first, the a pproach can be used accordingly for efficient system analysis of existing equipment with the aim to (retrospectively) equ ip them with sen sor app lications. This includes th e possibility to i ntegrate tech nical o r statistical analysis to point out crucial components for insta nce. Gain of knowledge incl udes the m ethodology fram ework t o analyse an d c hange e xisting sy stems. The framework a nd derived t ools su pport p roject engi neers i n asses sing as -is-situation an d designing sen sor ap plication an d co mplex rule system s for cond ition assessm ent i n a gu ided way. Moreover, the framework serves as an integrated possibility to transfer results of th e planning phase into the operative phase, acc ordingly wi th res pect t o t he f urther p rocesses. St ate-of-the-art al gorithms are i mplemented t o t est t heir performance in a real context.

Focusing on the case study, the main area of activity is the condition monitoring of systems and components using functional sen sors which work und er th e rough co nditions at seaports. Th e system h as to h ave a h igh av ailability an d reliability. Moreover, t he ongoing a pplication a nd im plementation of state-of-the-art m ethods to predict condition or failures an d its enh ancement will show , if th e co ndition can be assessed reliab ly. Thus, d ecision ru les can be implemented.

A general outcome of the work will create a v aluable contribution to th e paradigm shift of autonomous control in industrial pro cesses in cluding log istics b ecause t he ap proach en ables t he co nfiguration and operation of co ndition monitoring and therefore enables the systems to initiate further processes like maintenance or logistics actions based on particular decision rules. Research regarding the integration of information based on condition monitoring into planning and controlling of the logistical flows is an additional research topic.

5. REFERENCES

Byrne, G., Dornfeld, D., Inasaki, I., Ketteler, G., König, W. and Teti, R. (1995). Tool Condition Monitoring (TCM) -- The Status of Research and Industrial Application. *CIRP Annals - Manufacturing Technology*, *44*(2), 541–567.

Higgins, L. R., Mobley, K. and Wikoff, D. J. (2008). *Maintenance Engineering Handbook*: McGraw-Hill Education - Europe.

Jomdecha, C. and Prateepasen, A. (2009). Design of modified electromagnetic main-flux for steel wire rope inspection. *NDT & E International*, *42*(1), 77–83.

Kans, M. and Ingwald, A. (2008). Common database for cost-effective improvement of maintenance performance: Special Section on Advanced Modeling and Innovative Design of Supply Chain. *International Journal of Production Economics*, *113*(2), 734–747.

Lewandowski, M. and Echelmeyer, W. (2009). On the Application of Predictive Maintenance Strategies in Logistics: A Case-Study-Based Procedure Model. In *11th International Conference on The Modern Information Technology in the Innovation Processes of the Industrial Enterprises. Proceedings.* Bergamo.

Lewin, D. R. (1995). Predictive maintenance using PCA. *Control Engineering Practice*, *3*(3), 415–421.

Macián, V., Tormos, B., Olmeda, P. and Montoro, L. (2003). Analytical approach to wear rate determination for internal combustion engine condition monitoring based on oil analysis. *Tribology International*, *36*(10), 771–776.

Mobley, K. (2002). *An Introduction to Predictive Maintenance*: Society for Neuroscience.

Mobley, K. (2004). *Maintenance Fundamentals*: Butterworth-Heinemann.

Pham, H. T. and Yang, B.-S. (2010). Estimation and forecasting of machine health condition using ARMA/GARCH model. *Mechanical Systems and Signal Processing*, *24*(2), 546–558.

Scholz-Reiter, B., Görges, M. and Philipp, T. (2009). Autonomously controlled production systems--Influence of autonomous control level on logistic performance. *CIRP Annals - Manufacturing Technology*, *58*(1), 395–398.

Schuh, G., Kampker, A., Franzkoch, B. and Wemhöner, N. *Studie Intelligent Maintenance: Potenziale zustandsorientierter Instandhaltung*. Retrieved from http://www.ifm.de/ifmde/web/studie_ergebnisse.htm.

Swearingen, K. M. W. B. B., Gilbertson, D., Dunsdon, J. and Sykes, B. (2007). An Open System Architecture for Condition Based Maintenance Overview. In *2007 IEEE Aerospace Conference* (pp. 1–8). Big Sky, MT: I.E.E.E.Press.

DECISION MAKING MODEL FOR STRATEGIC ALLIANCE OF SERVICE CENTERS AND CONSOLIDATION TERMINALS IN EXPRESS COURIER SERVICES

KI Ho Chung[1], Hyun Jeung Ko[2], Friska Natalia Ferdinand[3], and Chang Seong Ko[4*]

[1]Department of Management Information Systems
Kyungsung University
314-79 Daeyeon-3dong, Nam-gu
Busan, Korea
e-mail: khchung@ks.ac.kr

[2]Department of Logistics
Kunsan National University
1170 Univ. Ave.
Gunsan, Korea
e-mail: hjko@kunsan.ac.kr

[3]Department of Indistrial & Management Engineering
Kyungsung University
314-79 Daeyeon-3dong, Nam-gu
Busan, Korea
e-mail: freezk_83@yahoo.com

[4]Department of Indistrial & Management Engineering
Kyungsung University
314-79 Daeyeon-3dong, Nam-gu
Busan, Korea
e-mail: csko@ks.ac.kr

1. INTRODUCTION

With the recent dramatic growth in delivery service market in Korea, the number of delivery service companies has also increased over twenties. However, more than 80% of the market is occupied by only four to five major companies such as Hyundai Logistics, Hanjin Express, Korea Logistics, CJ GLS, Korea Post. Except for a few well-known companies, however, most small and medium sized companies are in a difficult situation due to severe price competition in the market, difficulties in acquisition of delivery vehicles, and lack of country-wide terminals. Even though the market of express courier service in Korea has been expanding 20% or higher annually, which is much faster increase than in other industry areas, it becomes already saturated because large sized companies are entering the market through the acquisition of existing delivery service companies. Therefore, from the standpoint of national economy and the operational efficiency, the cooperative strategic alliances in the operation of facilities and delivery vehicles may be beneficiary especially for small and medium sized companies to overcome financial problems and improve the declining profitability by reducing or eliminating overlapped investments. Strategic alliance of cooperation creates economy of scale which leads to the reduction of operation cost. In addition, through efficient cooperation of service centers, participating companies may expect realizing an increase in net profit under a win-win situation. Providing better services to the customers by cooperatively utilizing their existing facilities, small and medium sized companies may efficiently compete with big-size companies by expanding their market share without further investment. The purpose of this study is to examine the feasibility of merge for under-utilized courier service centers for strategic alliance.

Therefore, in this paper, we propose a network design conceptual model, the strategic alliance model, to facilitate strategic alliance among small and medium express courier service companies. The model places participating companies in a win-win alliance relationship, and suggests how to increase the net profit of each company by harnessing their low demand and under-utilized service centers, and sharing consolidation terminals with available processing capacity. For this, we develop a nonlinear integer programming model and its solution procedure based on maxmin and maxsum criteria. And then to verify the applicability of the model to the real world problems, we presented a numerical example of the model using the data set collected from service centers and terminals of an express courier service company in Korea. The values for parameters are generated randomly for the purpose of simulation

2. LITERATURE REVIEW

To the best of our knowledge, not many studies directly related to the design of service network for express parcel delivery, have so far been undertaken. The study conducted by Cheung *et al.* (2001) was the first to examine a service network design problem encountering express couriers such as DHL Hong Kong. They proposed a hybrid optimization/simulation model that aimed to maximize service coverage and service reliability by adjusting cutoff time. Ko *et al.* (2010) suggested an approach to the reconfiguration of an express courier service network with respect to assignments of service centers to consolidation terminals and adjustments of their cutoff times. They proposed an integer programming model and a genetic algorithm-based solution procedure for allowing express couriers to maximize their incremental profit. On the other hand, there have been a few of researches on various types of strategic alliances in logistics. Many researchers, such as Chopra and Meindl (2004), Min (1996), and Simchi-Levi *et al.* (2003), pointed out that different companies owned and operated independently may benefit from the strategic alliance scheme which is conceptually similar to facility sharing and Cachon and Lariviere (1999) proposed a methodology for optimal capacity allocation. Recently, Chung *et al.* (2009) proposed a network design model for strategic alliances among express courier service companies by monopoly of service centers. Chung *et al.* (2010) extended their previous study to the problem of sharing consolidation terminals of each company. An integer programming model and its solution procedure based on a fuzzy set theoretic approach were developed. However, their study was performed under the assumption that the service centers only selected as candidates for strategic alliance are considered for sharing consolidation terminals.

3. PROBLEM STATEMENT

An express courier network generally consists of customer zones, service centers, and consolidation terminals. In the service center, express courier companies receive customer shipment request, pick up parcels, screen, sort, document, label, and consolidate them before transshipping them in bulk to a consolidation terminal. At the consolidation terminal, customer orders are consolidated into larger shipments, mixed and then loaded onto delivery trucks for local deliveries via service centers near consignees.

A strategic alliance model among the express courier companies for better utilization of their network can be classified into four models . Model (a) is a service center level collaboration model which is the weakest type of four alliance models. Each participating company operates its own terminals exclusively and collaborates in pick-up/delivery together at consolidated service centers. That is, a consolidated service center is not reallocated to each company. The objective of this model is to minimize the service center operation cost by reducing number of service centers in a scarce region. Model (b) is a weak terminal level collaboration model which is a medium level of alliance. The participating companies are able to share other company's terminals and reallocate service centers to the terminals. The goal of this model is to achieve a win-win strategy by improving the utilization rate of existing terminals of participating companies. Model (c) is a strong terminal level collaboration model which is the strongest level of alliance. This is a location/allocation model which simultaneously determines whether the existing terminals are still open or not and reallocation of service centers to the open terminals. Its objective is to minimize the total operation cost by reducing the number of terminals and reassigning the service centers to the terminals. Model (d) is an extended terminal level collaboration model which can be applied when participating companies have consolidated through Merger and Acquisition (M&A). This model determines whether the existing terminals are open or not with consideration of changing the terminal capacity. The goal of this model is to minimize the total operation cost and to maximize the utilization rate of terminals and service centers.

The strategic alliance model investigated in this paper is a hybrid model of the model (a) and (b) discussed above. In this model, companies participating in the alliance collaborate in pick-up/delivery together at the merged service center and share the capacity of consolidation terminals by reassigning all the service centers to the terminals. We assume that there are several regions in which each participating company has a service center but daily pick-up demand for each service center is very low. Now we call the regions 'merging regions'. Service centers and consolidation terminals are operated according to the following principles:

a) Multiple service centers can be open and all the other service centers are closed within a merging region after alliance.
b) The open service centers after alliance is also responsible for pick-up and delivery of all the amounts of other

company's service center s within the same merging region.

c) The current existing consolidation terminals can be still used even after strategic alliance.
d) Each company's service centers are divided into two types; Type I service centers belong to a merging region and are candidates for merging, and Type II service centers do not belong to any merging region.
e) The open service centers in each merging region as well as Type II service centers may be able to reassign to other company's consolidation terminal while satisfying the processing capacity of the terminal.

The problem of selecting one service center in each merging region and reassigning the open service centers as well as Type II service centers to the consolidation terminal is formulated as a nonlinear integer programming model and the solution procedure will be also developed with a numerical example in this study.

4. MATHEMATICAL MODEL

This section describes the integer programming model that has been designed to maximize the expected profit increase of each participating company. Suppose that there are n express courier companies, and that the location of terminals and service centers managed by each company are given. The underlying assumption is that there are m regions in which all n companies possess one service center independently with relatively small amount of shipment. Under the condition that the daily pick-up amount for the service centers in each region and the processing capacity of terminals per day are given, the problem is to maximize the profit of each company by selecting one among n service centers to be merged within each region and reallocating the merged service centers as well as normal service centers (Type II) to terminals subject to the processing capacity of all the terminals. It is assumed that all terminals operated by one company remain to be available to the other company after alliance.

The mathematical model has m objective functions which represent the net profits increases of m companies respectively. Since there exists a trade-off relationship between m objective functions, this paper proposes two different methods to derive the solution: maxmin criterion and maxsum criterion.

5. SOLUTION PROCEDURE

In this mathematical model (P), there are n objective functions representing the net profit increases of n companies respectively, and thus there exists a trade-off relationship among objective functions. In other words, it is inevitable to reduce the profits of other companies to maximize the profit of one company. Hence this paper proposes two methods using maxmin and maxsum criteria. First, a method that uses maxmin criterion is proposed to find compromised solution for the win-win situation of each company participating in the strategic alliance. In order to solve the problem using maxmin criterion the set of feasible solutions F(x) for the problem (P) and new variable α can be defined as follows:

F(x): Feasible solution set for problem (P), i.e., solution set satisfying constraints

$$\alpha = Min\{Z_1(x), Z_2(x), ..., Z_n(x)\}$$

By maximizing the value of α it is possible to derive the compromised solution which satisfies all n companies. The problem can be written as follows:

(P1)

$$\begin{aligned}
&\text{Max} \quad \alpha \\
&\text{s.t.} \quad z_1(x) \ge \alpha \\
&\qquad z_2(x) \ge \alpha \\
&\qquad \vdots \\
&\qquad z_n(x) \ge \alpha \\
&\qquad x \in F(x)
\end{aligned}$$

Second, this paper tries to solve the problem applying the maxsum criterion. The maxsum criterion is to maximize the sum of the net profits realized by all the companies within strategic alliance. It is intuitive, therefore, that the total net profits obtained from applying the maxsum criterion is greater than compared with maxmin criterion. However, it is desirable to distribute the realized total profits evenly among all the companies, which is not the case with maxsum criterion. The problem based on maxsum criterion may be formulated as (P2).

(P2)

$$\text{Max Z } = \sum_{i\in I}\sum_{j'\in J}(r_{ij'}D_i - f_{ij'} - \sum_{k\in T_{j'}} p_{j'k}a_{ij'k}D_i)x_{ij'} + \sum_{i\in I}\sum_{j'\in J}\sum_{k\in T_{j'}} p_{j'k}D_i y_{ij'k} + \sum_{j'\in J}\sum_{j\in J}\sum_{l\in S_j}\sum_{k\in T_{j'}} q_{jlk}d_{jl}^2 z_{jlk}$$

s.t. $x \in F(x)$

The objective function of (P2) is expressed as the sum of *n* objective functions in (P), which means the total sum of the net profits realized by *n* companies.

ACKNOWLEDGEMENTS

This resea rch was s upported by Basic Science Researc h Pro gram t hrough t he National R esearch Fo undation of Korea(NRF) funded by the Ministry of Education, Science Technology (2009-0089781).

6. REFERENCES

Chopra, S. and Meindl, P. (2004) *Supply Chain Management: Strategy, Planning and Operation*. Pearson Prentice Hall, NJ, USA.

Cachon, G.P. and Lar iviere, M.A. (1999). Cap acity Choi ce an d Allocation: St rategic B ehavior a nd S upply C hain Performance. *Management Science,* 45: 1091-1108.

Cheung, W., Leung, L.C. and Wong, Y.M. (2001). Strategic Service Network Design for DHL Hong Kong. *Interfaces*, 31: 1-14.

Chung, K.H., Rho, J.J. and Ko, C.S. (2009). Strategic Alliance Model with Regional Monopoly of Service Centers in Express Courier Services. *International Journal of Service and Operation Management*, 5: 774-786.

Chung, K.H., Ko, C.S., Hwang, Y.M and Rho, J.J. (2010). Network Design for Strategic Alliance in Express Courier Services: a Fuzzy Set Approach. *International Journal of Innovative Computing, Information and Control*, 9: 349-359.

Ko, H.J., Lee, H.J. and Ko, C.S. (2010) A Study on the Design of Express Courier Service Network Based on the Cut-off Time Adjustments. *International Journal of Computing, Information and Control*, 9: 335 – 347.

Min, H. (1996). Consolidation Terminal Location-Allocation and Consolidated Routing Problems. *Journal of Business Logistics*, 17: 235-263.

SEAPORT CONTAINER TERMINAL SERVICES ARE MOVING INLAND – CHALLENGES AND SOLUTIONS

Hans-Dietrich Haasis

ISL – Institute of Shipping Economics and Logistics
Universitaetsallee 11-13
28359 Bremen, Germany
haasis@isl.org

1. INTRODUCTION

The market environment in which container ports, shipping lines and port logistics operators are acting is rapidly and substantially changing due to a wide range of driving factors. For example, for Germany an increase in cargo transportation between 2004 and 2025 of 70 % is predicted. In line with this, especially seaport related hinterland transportation is concerned by more or less 130 %. Thus, the congestion at ports threatens progress and modal shift in the hinterland transportation network. Therefore, the role of intermodal inland terminals will gain much more importance in the future compared to their current use, sure enough after the medium term recovery of the worldwide economy. These terminals can serve as partners to seaport terminals, relieving them from capacity constraints on the one hand and offering value added services on the other. If all processes are properly controlled and managed, export containers can be sent to the seaports just before the vessel arrives; import containers can be brought to the designated inland hubs as soon as the containers are discharged and released. The attractiveness and economic success of a seaport will increasingly depend on the ability to integrate its business effectively into flexible supply chains reaching deep into the hinterland. Therewith the buffering and storage function of seaports can be well distributed to hinterland locations. These aspects get more and more in focus of different partners of the supply chains. In this paper related challenges in designing and operating these partnerships, decision support necessities and IT-based aspects are outlined, especially from the point of seaports of Germany.

2. OUTLINE OF THE PRESENT SITUATION

Regions are changing: World trade flows depend on markets, resources, logistics, production and innovation. Logistics ensure distributed production on a regional and on a global level. By this, new distributions of tasks and sites in international supply chain networks will occur. This has implications on hub functions and on new businesses, especially in port regions. Within the National Port Strategy of Germany, published in 2009 by the Federal Ministry of Transport, Building and Urban Development and seen as basis for the maritime policy for the next 10 years, measures for increasing the efficiency of the network of seaports and inland ports are described. A major role is foreseen for systemic supply chain orientated co-operations between container port-related companies as well as port regions, mainly between seaports and hinterland ports.

As an example the Hamburger Hafen und Logistik AG (HHLA) and Polzug are going to realize a network of transport relations on a regular basis between Hamburg and major European business centers, e.g. the hinterland terminal for combined transportation in Dabrowa Gornicza nearby Kattowitz. Also EUROGATE, Europeans number one of terminal operators, realizes since 2009 the establishment of an inland container terminal network, e.g. integrating Dortmund and Magdeburg.

All this is related to the fact that the strength of maritime regions is described by its relevance as partner in international supply chain networks. This relevance can be strengthened and improved by a process-overlapping system's optimization of production, transports, logistics, value adding services as well as regional development, and with this by a regional knowledge management. In the past, seaports had the function of bundling and coordinating trans-regional freight transportation as logistical nodes only. Nowadays seaports and maritime regions stay in the role of centres of logistics and value added services in intermodal global transport chains.

Continuing, seaports have an interface function in global transport chains. But seaport regions do not automatically offer the best business conditions for services, packing, further processing and dispersion of goods and products at the same time. The ongoing containerisation makes further dislocation of such value added activities in other regions possible. Logistics processes like containerisation and further processing of goods often can better be carried out efficiently in large consolidation centres in the hinterland than in the urbanized port areas in retaining small time slots. New nodes of international logistics systems arise in suburban areas or in peripheries of densely populated areas. That is the reason why the scientific perspective on seaports and port regions has changed during the last years towards a new paradigm considering ports as elements in value-driven chain systems. Therefore, there is need for action to strengthen the performance of the main gateways, there associated partners and benefits of coordinated strategy of maritime hub development by encouraging the development of vertical and horizontal not only regional but mainly functional clusters. However, clusters involving regional production and local sourcing provide an opportunity for reducing logistics costs and increasing welfare in port regions.

3. DECISION SUPPORT AND INFORMATION MANAGEMENT

As outlined, the importance of maritime transport and large-scale European ports will further increase in future due to the increase of global business. In line with this the European integrated maritime policy plays a major role for European development and integration. Having in mind the impressive forecasts on cargo flows and transshipments new solutions are required ensuring a secure and reliable transportation. Taking into account a systems perspective, a process-overlapping system's optimization of production, transports, logistics, value adding services as well as regional development is necessary. By this tools and instruments are required for organizational planning, decision support and financing on an operational and strategic level. According to the National Port Strategy especially cooperative solutions between companies as well as between regions or regional initiatives/clusters have to be regarded. For inland ports, and also for freight villages, this means to intensify the engagement in realizing seaport functions, e.g. warehousing and distribution of containers. The operations and functions within already existing corridors have to be analyzed as well as according to the results to be improved. New cooperative business models and new customer-orientated products have to be developed. Cooperative decentralized planning and controlling decision support systems have to be assigned and integrated in to the overall information system. Likewise, this means to undertake workflow analysis and business process reengineering. Decision support processes are related, amongst others, to the value-chain orientated assignment of processes and tasks, to location planning, to scheduling planning, to bottleneck capacity planning and to marketing management. Due to this, realizing the National Port Strategy, less technological than more operational innovation is requested. For these innovations besides methodical support provided by methods of operations research also measures for a team-orientated knowledge management related to human knowledge, process knowledge and structural knowledge, as well as for building-up of trust between the cooperation partners are required. In line with balancing of the challenges, the use of an adequate cost and risk sharing model is essential.

The feasibility and cost-efficiency of these port co-operations and of cause also of the terminal driven solutions depend on an information management between terminal operators, logistics service providers and providers of infrastructures. There is a need for a harmonisation or integration of existing regional port information systems, business information systems and European information system developments. Thus, the balance is still open between the top down D2D SCM business approach and the bottom up infrastructure approach concerning the traffic layer, the business process layer and the information layer. The objective is to realize an information architecture which is able to handle the above described situation (information layer) and business models according to the requirements of industry, retailers, forwarders and ports (business process layer). By this, a new corridor orientated intermodal transportation strategy in maritime business can become reality (traffic layer).

4. CHALLENGES AND OPPORTUNITIES

The attractiveness and economic success of a seaport will increasingly depend on the ability to integrate its business effectively into flexible supply chains reaching deep into the hinterland. Major success factors are reachability and accessibility, availability of transshipment capacities and storage capacities, integration into a network of hinterland relations, the quality of value added services as well as the customer-related transshipment costs. As mentioned, challenges cover the increased use of transport infrastructure, the increase in transshipment, the closer cooperation with market regions and the assignment of responsibilities and tasks between the seaport and the inland port. Opportunities are the related economic growth, the opening up of new markets, the realization of a green and reliable intermodal solution by bundling of transports, the sharing of capacities, costs and risks as well as the improved performance profile of the cooperation partners. By this extended gate solution, finally the terminal operator and the seaport operate closer to the customer. Based on this integrative perspective both, the terminal operations as well as the inbound and outbound transportation flows can be optimized.

5. REFERENCES

Bartowiak, J., Schreiner, U., Siestrup, G., Haasis, H.-D. and Plöger, M. (2009). A Flexible Communication Concept for Integrated Supply Chain Planning Concerning Aspects of Trust. *Proceedings of the 14th International Symposium on Logistics*, 252-257.

Bundesministerium für Verkehr, Bau und Stadtentwicklung (ed.). *Nationales Hafenkonzept für die See- und Binnenhäfen*, Berlin, 2009.

De Langen, P. (2008). Ensuring Hinterland Access: The Role of Port Authorities, *JTRC OECD/ITF Discussion Paper 2008-11*.

Elbert, R., Haasis, H.-D., Schönberger, R. and Landwehr, T. (2009). Adapting Dynamic Logistics Processes and Networks – Advantages through Regional Logistics Clusters, *Proceedings LogDynamics International Conference*, 2009.

Giuliano, G. and O'Brien, T. (2008). Responding to Increasing Port-related Freight Volumes: Lessons from Los Angeles / Long Beach and Other US Ports and Hinterlands, *JTRC OECD/ITF Discussion Paper 2008-12*.

Haasis, H.-D. (2005). Mass Customization in International Logistics. Blecker, T. and Fiedrich, G. (eds.). *Mass Customization. Concepts – Tools – Realization*, Berlin, 189-193.

Haasis, H.-D., Wildebrand, H., Plöger, M. and Zimmermann, F. (2009). Traffic reduction through intelligent control of freight service on the basis of autonomous multi-agent transport coordination, *Proceedings of the 13th EWGT Conference*, Padua, 2009.

Haasis, H.-D., Möllenstädt, O. (2006). Strategic Tools for the Sustainable Development of Maritime Regions. Haasis, H.-D., Kopfer, H. and Schönberger, J. (eds). *Operations Research Proceedings 2005*, Berlin, 123–128.

Haralambides, H. E. (2007). Structure and Operations in Liner Shipping Industry, Button, K. J. and Hensher, D. A. (eds.). *Handbook of Transport Modeling*. 2007.

Herer, Y. T. and Tzur, M. (2001). The Dynamic Transshipment Problem. *Naval Research Logistics*, 48: 386-408.

Mangan, J., Lalwani, C. and Fynes, B. (2008). Port-centric logistics. *International Journal of Logistics Management*, 19: 29-41.

Moyaux, T., Chaib-draa, B. and d'Amours, S. (2006). Supply Chain Management and Multiagent Systems: An Overview. Müller, J. P. and Chaib-draa, B. (Hrsg.). *Multiagent-Based Supply Chain Management*, Berlin, 1-27.

Notteboom, T. (2008). The relationship between seaports and the intermodal hinterland in light of global supply chains. European challenges, *JTRC OECD/ITF Discussion Paper 2008-10*.

Notteboom, T., Ducruet, C. and Langen, P. de (eds.) (2009). *Ports in Proximity. Competition and coordination among adjacent seaports*, Hampshire.

Notteboom, T. and Rodrigue, J. P. (2005). Port regionalization: towards a new phase in port development, *Maritime Policy and Management*, 32: 297-313.

Pawellek, G. and Schönknecht, A. (2007). Maritime Logistik – Innovationsstrategien, Lösungsansätze und Potenziale. *Jahrbuch Logistik*, 87-89.

Robinson, R. (2002). Ports as elements in value-driven chain systems: the new paradigm. *Maritime Policy and Management*, 29: 241-255.

Rodrigue, J.-P., Comtois, C. and Slack, B. (2009). *The Geography of Transport Systems*. New York.

Stopford, M. (2008). *Maritime Economics*, Oxford.

Talley, W. K. (2009). *Port Economics*, Oxford.

Wang, J., Oliver, D., Notteboom, T. and Slack, B. (eds.) (2007). *Ports, Cities, and Global Supply Chains*, Hampshire.

Session C4: Maritime 1

·**Day1: Sep. 15 (Wed.)**

·**Time: 15:00 - 16:20**

·**Chair: Won Young Yun**

·**Room: Grand Ballroom, 5F**

LOGMS

The 1st International Conference on Logistics and Maritime Systems

September 15th–17th, 2010 Busan, Korea

Lin

A STUDY ON SELECTION CRITERIA OF GREEN PORT AND EVALUATION OF MAJOR KOREAN PORT

Ji-Yeong Pak[1] and Gi-Tae Yeo[1*]

[1]Graduate School of logistics
Incheon University
408-840
Incheon, Republic of Korea
e-mail: assambleuse@hanmail.net, ktyeo@incheon.ac.kr

Abstract: The importance of environmental concerns in managing a port is being emphasizing. And so this trend toward eco-friendly port will affect port competitiveness. A lot of ports in the world have made a start on suitable development of a port. For this reason, the main purpose of this study draws out the criteria of green port that is available and possible in practice. In addition, this study decides the ranking among Korean five trade ports (In-cheon, Bu-san, Gwang-yang, Pyeong-taek, Ul-san) with selected the green port criteria. And, Fuzzy methodology is chosen for evaluating green ports for reason of taking measurements its complexity and ambiguity.

1. INTRODUCTION

World Commission on Environment and Development (WCED) has tried to excite green issue of 'sustainable development' in politics and economics all over the world. And WCED commented that transportation industry play a great role in environmental damage (Button Henshe, 2001).

Green Logistics means ease the burden on environment and achieve economical rationality through improvement of transportation efficiency. Moreover, promoting with green Logistics can help improve business competition by cutting down on distribution costs as well as reduce CO_2 emission. In other words, the activities of green logistics imply to build an efficient logistics systems which can reduce environmental burden causing from the entire logistics environment, including management and disposal of waste products.

In the case of a port, lots of main ports in the world have started to work to preserve the environment of a port in many parts. For examples, Tacoma port in USA progresses a broader definition of “environment" - one that includes the people, businesses, culture and history of a place, along with its native species and natural systems. Meanwhile, in case of Japan, port environmental policies are going into effect more thoroughly than any other countries. The ministry of transport of Japan promotes new policy that a port is pleasant and enjoyable space in harmony with environment rather than passive environmental policy reducing negative effects according to the use and development of a port. And the aim of port of Rotterdam in Netherlands is to promote economic activities inside a port and is to maintain a clean as well as safe environment at the same time. Therefore, the authority of Rotterdam port has a plan to improve both its performance and that of the port when it comes to sustainability

In terms of these green port trends, Korea has to also take a step for keeping up with developed green system. Korea Thus, this study proposes eco-friendly port factors and shows the green criteria for a green port fact or analysis. Taking a step forward, this study decides the ranking among Korean five trade ports (Incheon, Busan, Gwan-yang, Pyeongtaek, Ulsan) with selected the green criteria.

2. LITERATURE REVIEW

Domestic and foreign researches on an eco-friendly port and port pollution are not enough for now but it will be getting more increase because of the importance of the problem. Therefore, the mainstream of these domestic researches on an eco-friendly port has been treated aspects of legislation and institution (Song and Han, 2007). Choi *et al* (2000) stressed introducing of total amount of port pollutant regulation and management system according to pollution source after investigation about wastes, water quality, and pollution from ship in Ma-san Port. Park (2004) suggested the real condition of environment management by a port by applying Self Diagnosis Methodology that was made of scholars in advanced countries. And Korea Maritime Institute (2005) introduced the trend of correspondence of United Nation and major countries about air pollution control and management in a port.

Meanwhile, Gallagher *et al.* (2003) shows how international trade is influencing air pollutions in the shipping sector of U.S.A through the method of economic costs estimation for air pollutions. Heand Morrison(2001) studied the result of Kembla port from 1975-1995, on water and sediment quality, benthic organisms and pollution levels, and comment on trends in the environmental health of the port and pollution reduction practices of local heavy industries.

Only a few foreign studies that have been published by the experts who are in the field of environment and there is not much systematic analysis from port logistics experts.

Table1. Variables for selecting green port by literature review

Variables Description	
Resources recycling inside a port	-Reduction of waste of resources
Technical development of industry and the ocean waste disposal	-A method of eco-friendly construction
Development of Breakwater system of waterfront revitalization	-Creation of a pleasant life environment
Dredging sand recycling	-Utilization as a reclamation material after solidification -Using at artificial wetlands or beach face development
Creation of artificial sandbar and wetland	-A recovery of natural purification of an ecosystem
Establishing Strategic Plan for Port environment and Introduction of environmental impact assessment	-Exploring every avenue about reducing environmental effects by predicting harmful effects on port environment.
Use of alternative fuels	-Solar energy, Wind energy, tidal energy and so on
Introduction of port environment management system	-Working in close cooperation with companies and administrative agencies.
Port facilities and equipment improvement	-AMP(Alternative Maritime Power) -DPF(Diesel Particulate Filter trap) - RTGC
Incentives of pollution reduction	-Incentives to shipping companies and stevedores with eco-friendly equipment.
Use of Renewable-energy	-CNG(Compressed Natural Gas), Bio-Diesel, hydrogen fuel
Modal shift	-Repair and maintenance a harbor-side road and Introduction of new transportation modes.
Construction method of noise reduction	-Use of construction methods for noise and vibration control -Use of noise reduction equipment
Efficient construction plan	-Eliminating unnecessary procedures
Expansion of the coastal region facilities (prevention of ocean pollution caused by land activities)	-Expansion and improve the sewage disposal plants, sewage landfills, waste water disposal plants, and so on

Source: Pak *et al.* (2009)

3. EMPIRICAL ANALYSIS

3.1 Application of Factor Analysis

This study uses a factor analysis of SPSS to examine the factors that influence on selecting green port criteria. Factor analysis is a statistical approach that can be used to analyze interrelationships among a large number of variables and to explain these variables in terms of their common underlying dimensions factors. The statistical approach involving finding a way of condensing the information contained in a number of original variables into a smaller set of dimensions with a minimum loss of information (Hair *et al*, 1992). In other words, factor analysis is a method of data reduction.

Before examining factor analysis, it is useful to draw out the factor of eco-friendly port. These factors are drawn out from literature review. And the data reduction of these factors is carried out by surveys of experts who are engaged in port public corporations, shipping companies and global terminal operators. The survey obtained 15 selected criteria by Literature Review and was executed by Likert Scaling.

The Factor analysis is carried out with the criteria in Table 1 and extracts five factors.

Table2. Rotated Component Matrix

	Component				
	1	2	3	4	5
Use of alternative fuels	**.742**	.129	-.082	.051	-.084
Incentives of pollution reduction	**.725**	-.005	-.276	.365	.119
Renewable-energy using	**.644**	.244	.057	-.090	.384
Dredging sand recycling	**.553**	.090	.425	.026	.093
Port facilities and equipment improvement	.168	**.796**	-.109	.065	-.072
Development of Breakwater system of waterfront revitalization	.041	**.673**	.305	.027	.116
Construction method of noise reduction	.283	**.521**	.229	.178	.237
Technical development of industry and the ocean waste disposal	-.150	.158	**.818**	.014	-.063
Resources recycling inside a port	.040	.030	**.795**	.262	.028
Introduction of port environment management system	-.011	.474	-.151	**.670**	-.100
Expansion of the coastal region facilities	.042	.189	.306	**.657**	-.005
Efficient construction plan	.240	-.172	.248	**.651**	.152
Modal shift	.134	-.055	-.074	-.062	**.844**
Establishing Strategic Plan for port environment and Introduction of environmental impact assessment	.008	.071	-.044	.519	**.572**
Creation of artificial sandbar and wetland	.043	.459	.168	.128	**.565**

Source: Pak *et al.* (2009)

From the above factor analysis result, five factors may be drawn about the green port criteria: (1) Ease the environmental burden (2) Environment-friendly method & technology development of construction (3) Utilization of resources & waste inside a port (4) Efficient planning & management of port operation (5) Port redevelopment as waterfront belt.(Pak *et al.* ,2009)

3.2 Fuzzy Method

3.2.1 Fuzzy Set

Fuzzy sets theory which accepts linguistic ambiguity was introduced by L.A.Zadeh (1965) as a widening of the classical concept of set. Fuzzy theory has specific features in comparison with an existing theory. Fuzzy sets are sets that have degrees of membership of elements. The membership of elements in a classical set is assessed in binary terms according to bivalent condition an element either belongs or does not belong to the set. Conversely, fuzzy set theory is allowed to the gradual assessment of the membership of elements in a set. This is told with the aid of a membership function valued in the real unit interval [0, 1]. Fuzzy sets generalize classical sets, because the indicator functions of classical sets are special cases of the membership functions of fuzzy sets, if the latter only have values 0 or 1. Classical bivalent sets are in fuzzy set theory commonly called crisp sets. Triangle fuzzy number $\overline{A}$ consists of three parameters (a_1, a_2, a_3) and the $\overline{A}$ 'membership function $\mu A(\chi)$ is defined as seen in Equation1.

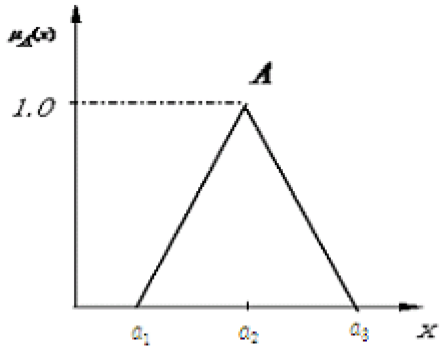


Figure1. Function of Triangle Fuzzy Number

$$\mu A(\chi) = \begin{cases} 0, & \chi \prec a_1 \\ \dfrac{\chi - a_1}{a_2 - a_3}, & a_1 \le \chi \le a_2 \\ \dfrac{a_3 - \chi}{a_3 - a_2}, & a_2 \le \chi \le a_3 \\ 0, & \chi \succ a_3 \end{cases} \qquad (1)$$

Ith triangle fuzzy number among membership function 'n' is defined as in Equation2.

$$\overline{A} = (a_1^{(i)}, a_2^{(i)}, a_3^{(i)}, \quad \mathrm{i} = 1,2,3\cdots\cdots, \mathrm{n} \qquad (2)$$

And, fuzzification $\overline{A}$ is defined as follows.

$$\overline{A} = A_{ave} = \frac{\overline{A_1} + \overline{A_2} + \cdots + \overline{A_n}}{n} = \frac{(\sum_{i=1}^{n} a_1^{(i)}, \sum_{i=1}^{n} a_2^{(i)}, \sum_{i=1}^{n} a_3^{(i)}}{n} = (a_1, a_2, a_3) \qquad (3)$$

The strength of fuzzy of Zadeh has been emphasized in the way that it can be quantitative linguistic variable. The linguistic term does not mean number but variable measured as form of word or sentence. The linguistic term is able to denote a fuzzy set. Figure3 is shown triangle membership function about the values of linguistic measures [Very bad - bad - medium - good - very good]

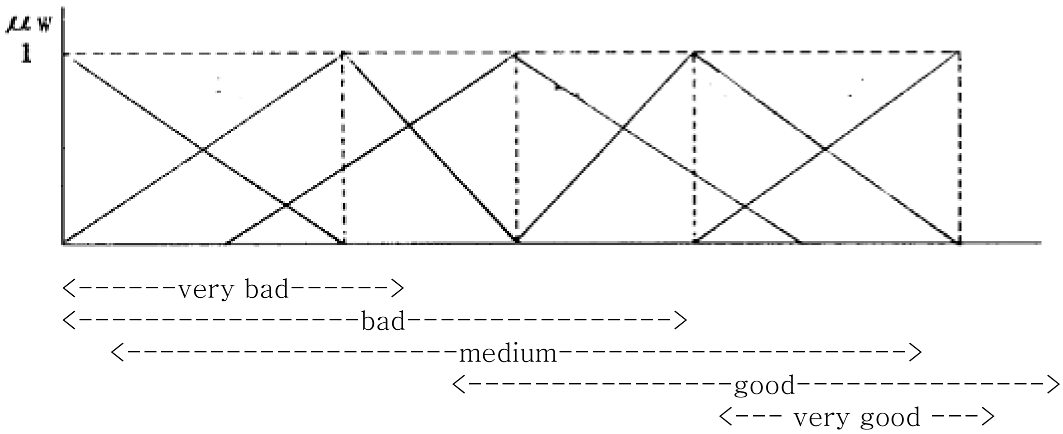


Figure2. Linguistic term of triangle fuzzy number

After above process, it is necessary triangle fuzzy number to convert as defuzzification. The aim of defuzzification is to convert fuzzy set the entire result obtained in previous step into an actual number that, somewhat, best represents fuzzy set. The most common defuzzification methods are Center of gravity. Center of gravity is one of the most commonly used defuzzification techniques. This method resolves the centre of the area of the combined membership functions.

$$y^* = \frac{\sum \mu(yi)xi}{\sum \mu \ \ yi} \tag{4}$$

In the latter part of this study, y* represents as G-factor and G-factor is defined value by center of gravity.

3.2.2 <u>Application of Fuzzy Set</u>

Fuzzy method is chosen for evaluating green ports for reason of taking measurements its complexity and ambiguity. Actually, it is not easy to evaluate its qualitative factors of green indexes of ports. Collecting the quantitative data about qualitative factors is limited. For all the reasons given previously, this study is evaluating the ports by fuzzifying the result from experts' survey. Especially, in the light of port evaluation, it is very difficult to determine the ranking with simple quantitative data given the paucity of material on specific port's quantitative datum.

As the case stands, experts' survey is a proper consideration by evaluating port's a certain degree of greening. Survey was conducted from October 05, 2009 to October 10, 2009. The surveyees of the survey were port public corporations, shipping companies and global terminal operators.

The result of fuzzy method based on experts' survey is below.

<Table3> Port ranking with factors

G-factor(Ranking)					
	Incheon	Busan	Gwangyang	Pyeongtaek	Ulsan
Factor1	0.457(3)	**0.528(1)**	0.487(2)	0.398(5)	0.383(6)
Factor2	0.500(3)	**0.631(1)**	0.560(2)	0.455(4)	0.455(4)
Factor3	0.357(3)	0.444(2)	**0.469(1)**	0.383(5)	0.368(4)
Factor4	0.487(3)	**0.601(1)**	0.573(2)	0.485(4)	0.442(5)
Factor5	0.454(3)	**0.550(1)**	0.509(2)	0.426(4)	0.406(5)
Overall Ranking	3	**1**	2	4	5

Busan took the No. 1 ranking the result of defuzzification by experts' survey about factor1 (ease the environmental burden). And Gwnag-yang held the second rank by 0.041.

About factor2 (environment-friendly method & technology development of construction), Busan ranked first again this time and Gwang-yang is at the back of Busan. Busan is changing diesel which is a power source of RTGC into electricity and thus it is projected to decrease 202 tons per year in nitrogen compound, 9tons in diesel dust, 18.2 billion won in operating expense and 50% in engine failure. In addition, new docks in Busan port is progressing facility construction like underground tracks for electrical power, manhole. It seems to be one of reasons to be selected as a first ranker in factor 2.

The third factor is utilization of resources & waste inside a port. Gwang-yang ranked first on the third factor.

Busan took the No.1 ranking again the result of defuzzification by experts' survey about factor4 (efficient planning & management of port operation). Lastly, the fifth factor is port redevelopment as waterfront belt. And Busan ranked first on the fifth factor. A port waterfront means introducing factors and facilities for recreation, physical training and demonstration function according to the progress of urbanization. Overall, Busan ranks top in four factors of five factors.

4. CONCLUSION

This study gives a new light in recognizing the role of the ports and this light has not seen before in this field. On the one hand, it is a first attempt in driving green criteria and it tries to rank ports of Korea by selecting the green port criteria. The present level of a port concerns the extent of green that can be revealed with these 5 factors. These five factors are as follow: (1)Ease the environmental burden (2) Environment-friendly method& technology development of construction (3)Utilization of resources& waste inside a port (4)Efficient planning & management of port operation (5)Port redevelopment as waterfront belt. With these criteria, Busan takes the No.1 ranking the result of fuzzy by experts' survey. About factor1 (Ease the environmental burden), factor2 (Environment-friendly method and technology development of construction), factor4 (Efficient planning & management of port operation), Factor5 (Port redevelopment as waterfront belt). And Gwang-yang ranked first on the third factor (Utilization of resources & waste inside a port).

Evaluating trading ports in Korea can provide the new course understanding the new trend and the role of a port in the 21st century.

5. REFERENCES

A, V, Patel and B, M, Mohan (2002). Some numerical aspects of center of area defuzzification method. *Fuzzy Sets and Systems 132,* pp.401-409.

Bailey, D and Solomon, G (2004). Pollution prevention at ports: Clearing the air. *Environmental Impact Assessment Review.*

Choi, D, H and Choi, J, H and Mok, J, Y and Park, Y, U and Yun, S, S (2000). Major port pollution research and management plan. *KMI*

Jeong, B, H (2009). Green port management policy directions in the green growth era - The case of Gwangyang port in Republic of Korea. *Journal of Korea Port Economic Association*, Vol. 25(3) p.361-384.

Kevin, Gallagher and Robin Taylor (2003). International trade and air pollution: The economic costs of air emissions from waterborne commerce vessels in the United States. *Global Development and Environment Institute Working Paper No.03-08.*

Lee, H, S and Han, D, H and Choi, Y, J. A study on service quality measurement using triangle fuzzy number(TFN) and analytic hierarchy process(AHP). *Korea Industrial and Systems Engineering*

Lee, S, H and Yun, D, G (2004). Using fuzzy numbers to evaluate service quality (FR-SERVQUAL). *Korea Industrial and Systems engineering.*

Pak, J, Y and Yeo, G, T and Coo, B, M (2009). A study on selecting green indexes of a sea port. *International Journal of Navigation and Port research Vol.33, No.9*, pp.653-657

R. M. Dabra and A. Ronzaand j. Casal and T. A. Stojanovic and C. Wooldridge(2004). The self diagnosis method a new methodology to assess environmental management in sea ports. *Marine pollution Bulletin* 48, pp. 420-428.

Song, G, E and Han, C, H (2007). A Study on the Strategies for the Reduction of Port Pollution. *Journal of Korea Port Economic Association*, Vol. 23(1), pp.95~113.

Zhijia he and R.J. Morrisin (2001). Changes in the marine environment of Port Kembla harbour, NSW, Australia, 1975-1995:A Review. *Marine Pollution Bulletin* Vol. 42, No.3, pp.193-201.

HAVENBEDRIJF ROTTERDAM N.V. SUSTAINABILITY REPORT 2008 (Summary of 2008 CSR annual report). www.portofratterdam.com

Port of Tacoma. www.portoftacoma.com

INNOVATIVE TRANSPORT SYSTEM FOR SEAPORT TERMINALS - ROPEWAY FOR MARINE CONTAINERS

Ann-Kathrin Pallasch[1*], Jan Heitkötter[1] and Bernd Scholz-Reiter[1]

[1]BIBA Bremer Institut für Produktion und Logistik GmbH
at the University of Bremen
Hochschulring 20
Bremen, Germany
e-mail: pal@biba.uni-bremen.de

Abstract: Sea ports have experienced tremendous growth rates that have lead to a shortage of free storage capacity. Due to the world economic crises in 2008, the container turnover rates broke down but are expected to recover soon and reach or even exceed pre-recession level (cf. Maaß *et al*., 2010). This requires enlarging storage capacities for standard marine containers and consequently enlarging storage areas in seaport terminals. In many cases structures around seaports like roads, rails, or buildings inhibit their expansion. To establish an efficient raise of storage capacities and their connection to the home terminal, an innovative, bidirectional, fully automated infrastructural module has to be developed: the "Ropeway for Marine Containers". This technology is supposed to be usable for various scenarios. Thus, this transport module will allow an efficient, ecological transport between different locations.

1. INTRODUCTION

Until the world economic crises in 2008, the maritime container transport had shown constant growth rates (as shown in Figure 1). In 2009 and the beginning of this year, the world maritime container transport started to rise again. Many experts are sure, that worldwide trade will recover soon and reach or even overtake the pre-recession level. To efficiently meet the requirements of a high level of world maritime container transportation, certain innovative technologies for seaports and the surrounding processes have to be developed.

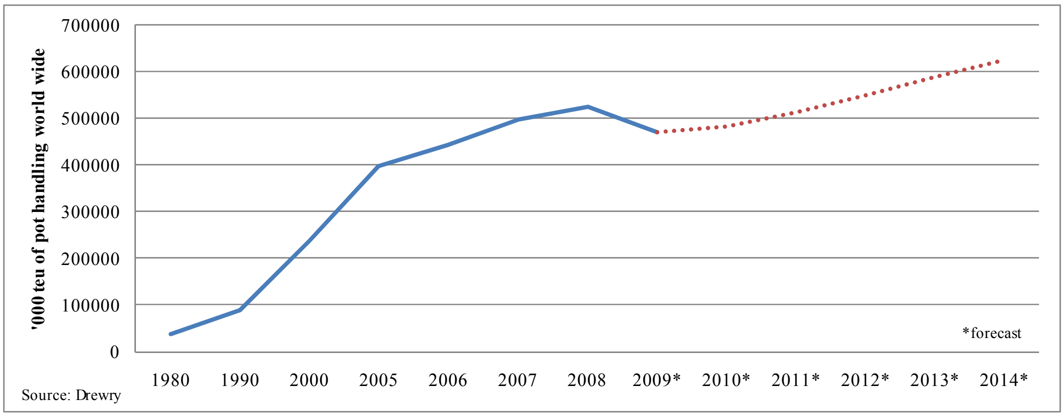


Figure 1. Container Activity Worldwide: 1980-2009 (Gardiner, 2009)

In order to meet the demands of high level container activities worldwide, seaports had to expand their storage capacities steadily. Many seaports are a result of structures that have evolved over a long time. Hence, they are connected to the surrounding environment by different kinds of transportation, such as, road, water or rail. Thus, there are several physical barriers, like water on the one side of the terminal and existing infrastructure like buildings, roads or tracks on the other side, that inhibit the enlargement of storage capacities for standard marine containers. With these aspects in mind, a research project at BIBA (The Bremen Institute for Production and Logistics at the University of Bremen) is studying the technical and economical feasibility of an innovative infrastructure module that connects separated storage areas for marine containers: the ropeway for marine containers. This research project is funded by the „Federal Ministry of Economics and Technology".

The project focuses on a special scenario, settled on a large seaport terminal in Hamburg (as seen in Figure 2).

After the feasibility has been confirmed, the next goal will be to design and build the “Ropeway for Containers” in a second phase of the project.

Figure 2. Container terminal in Hamburg (Port of Hamburg, 2009)

Regarding this setting, the enlargement of the terminals’ storage areas is inhibited by different physical barriers. Under these conditions, the operating company has established a new depot area, very close to the home terminal (as seen in Figure 3. and Figure 4.). At present these separated storage areas (depot and home terminal) are connected by special freight vehicles, so-called “multitrailers”. This system consists of a towing vehicle and a variable number of trailers. A multitrailer can transport up to 6 TEU (TEU: twenty feet equivalent unit - one 40ft. container corresponds 2 TEU) per trip. Thus, their total length is approximately 45 meters which requires a special licence that permits passing public roads with one multitrailer each time. Several multitrailers are provided by the terminal operator to maintain an average output of 40 TEU per hour (1-shift operation), 360 days a year.

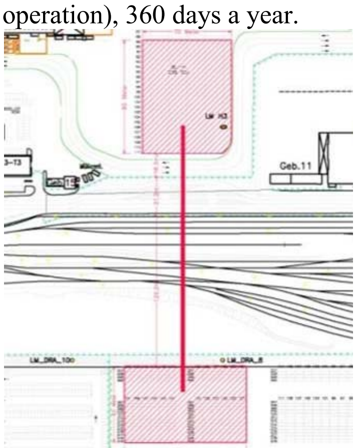


Figure 3. Overview (Source: project partner)

The strait-line distance between these two areas is around 200 meters; the multitrailers have to cover a distance of around 4 kilometres. A public road, a terminal road and tracks would have to be crossed in a direct connection. Furthermore, the system used leads to high fixed costs and CO_2 emissions. Consequently, an innovative infrastructure module that allows a direct connection is needed to provide an efficient increase of the storage capacities.

Figure 4. Birds-eye-view on the terminal (Source: project partner)

Wherever physical barriers have to be crossed to expand seaport terminals' storage capacities, there has to be a special focus on the initiated traffic between these areas – traffic that connects the added area with the home terminal. Today's technologies have not led to an adequate solution. The use of the ropeway for containers shall lead to benefits like simplification of the connection between storage areas, traffic reduction as well as reduction of CO_2 emissions and minimisation of short distance transport costs.

2. DEMANDS ON THE TECHNICAL SYSTEM

Ropeways are usually used in the mountains to transport persons or goods from a valley station to the top of a mountain. In order to implement a ropeway in the area of a seaport, special environmental demands have to be considered, for example, contrary to the situation in the mountains, the ground at seaports is mostly sandy. Furthermore, seaports are built very close to the water line, which means that the area is flat. Hence, there will be no level difference between the two stations. As a result, the containers always have to be lifted if there are any barriers along the route.

Investigations of the actual state and discussions with the terminal operator have lead to functional and non-functional demands. These demands have been summarized in a product concept catalogue and are presented in the following.

Functional demands:

1. The transport system has to run fully automatically. That means that there are only two interfaces with the system: a freight vehicle like a Straddle Carrier (SC) or Reach Stacker (RS) has to deliver a container to the system and a transported container has to be lifted and forwarded by e.g. a SC or RS.
2. To uncouple the cycle time of the ropeway system from the cycle time of the freight vehicles, the system has to be fitted with a buffer store on each side.
3. The system should run 24/7, 360 days a year.
4. The ropeway for marine containers should be able to transport 20 ft. and 40 ft. standard containers.
5. The containers may have a maximum load of about 20 tons. The feasibility study shows that more than 80% of the containers have less weight than 20 tons. Thus, the payload of the ropeway should be around 20 tons; gross weight should be around 25 tons.
6. Distance that has to be covered: 200m.
7. Transportation of up to 60 containers per hour should be possible.
8. Two independent tracks should be implemented to get reliable redundancy of the system.
9. The tracks shall be usable in both directions so that if there are transportation peaks in one direction, both tracks can be operated simultaneously in the same direction.

The non-functional demands on the system are:

1. In terms of economic efficiency the system should feature an adequate cycle time. It depends on: distance between stations of reception and release, kind of barriers / area which the ropeway bridges, ropeway technology (maximum speed, time of loading / unloading, further specifications).

2. The system should be ecologically efficient. This has to be considered due to its conception. It has to be ecologically acceptable in relation to its construction, its lifetime and the recycling of the used-up ropeway.
3. The operational availability of the system has to be around 98% (expressed in hours per year, excluding maintenance).
4. At seaports, soil conditions are very different regarding common soil conditions of ropeways in the mountains (e.g. sandy ground in areas that are located close to water lines, e.g. seaports).
5. The system has to meet local and general safety and security regulations and laws.
6. There are a lot of building and machinery regulations at seaports e.g. regarding the allowed amount of noise. These local regulations have to be taken in account during construction of the ropeway.
7. In addition, facilities at seaports underlie very rough conditions caused by weather, salty air, imprecise works with heavy loads, etc. Therefore, everything has to be built to hold up to the prevailing conditions and satisfy special requirements regarding e.g. corrosion.
8. The initial system should possess a high adaptability for different global conditions (other conditions of subsoil, different length of transportation,...).
9. The integration of the new transport module into the existing infrastructure should be easy.

The specifications of the technical system and its subsystems have to meet all requirements. Later on, the different concepts for a ropeway for marine containers will be examined regarding these requirements.

3. DIFFERENT SOLUTIONS AND THEIR FEASIBILITY

The general, theoretical processes in action and functions of the ropeway for marine containers are shown in the following figure (Figure 5). At each station freight vehicles like Straddle Carriers or Reach Stacker may transfer the containers to the buffer of the ropeway (RECEPTION). Thereafter, the buffer stores and forwards the container to the next available buffer place, by using a defined number of separated buffer places. Finally, the marine container gets transferred to the transfer station fully automatically. The ropeway performs the transport to the opposite side. Having been transferred to the buffer on the other side of the physical barrier and forwarded to the release position, the detachment and further transport of the container can be executed by any container transport vehicle.

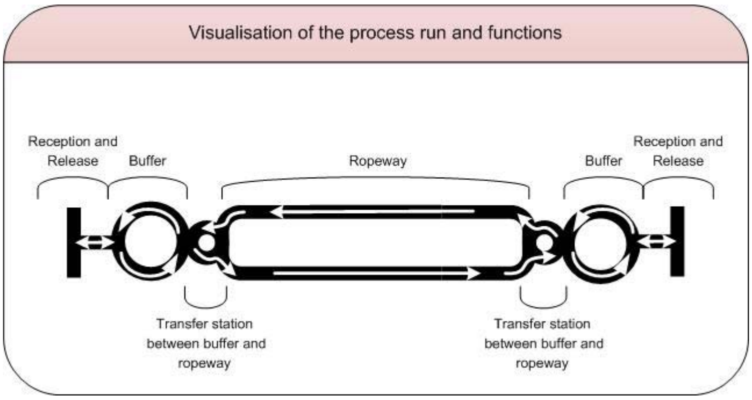


Figure 5. Process run (Source: Authors)

The existing technical systems of ropeways can be divided into two main methods of construction (Schweizerischer Bundesrat, 2007). In the first type, which is called the “funicular ropeway”, the gondolas are designed as wagons that move along a railway. The towing rope is only used to pull the gondolas in the desired direction. The second general type is called “aerial ropeway”. This system uses different cables to gather the load and to pull the gondolas. As described in the following sections, the chosen technical concept of the ropeway has a great influence on efficiency in each specific case. Due to those different technical types of ropeways there are also different possibilities to use them for the transportation of marine containers.

3.1 Funicular Ropeway

As shown in the following figure (Figure 6.),a funicular ropeway is designed as a train or wagon that runs on tracks that are built on a substructure and end in a lower and an upper station. In contrast to a conventional train, they are not self-powered but are pulled by a towing rope. The rope is attached to each wagon and is moved by a cable winch that is installed in the stations. The funicular ropeway system always needs a solid substructure, which provides a solid basis for the tracks.

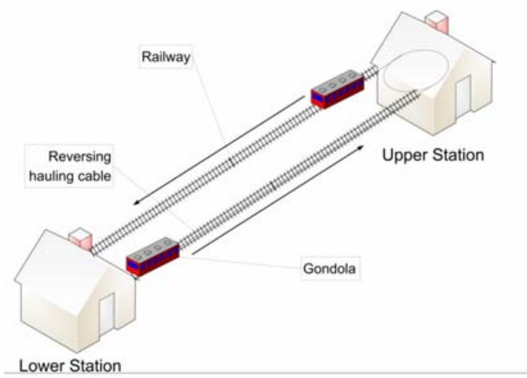


Figure 6. Funicular ropeway (Source: Authors)

One option for constructing a substructure to get over obstacles is to build concrete structures like bridges. It has to be taken in account, that in addition to the maintenance of the ropeway, that substructure also has to be maintained.

Funicular ropeways are normally designed as a twin system with a circular haulage cable, to provide optimal energy efficiency. This system takes advantage of two equal rail tracks. Whilst one wagon is pulled from the valley station to the upper station the wagon on the other track is moving in the other direction. In doing so, the potential energy of one wagon can be used to pull the second; only a small amount of additional energy has to be provided.

The cycle time of funicular ropeways is relatively high (as compared to circulating ropeways) but they are constructed to transport high payloads. To implement the substructure, bridge-like concrete structures are needed. Large obstacles that have to be crossed, barriers like high buildings or wide spaces like rivers, complicate the installation of the substructure. As it is only possible to transport one wagon at a time in each direction, the cycle time of funicular ropeways depends on their length; but there might be installed several ropeways linked in a chain to increase the cycle speed regarding medium or long distance. Additionally, the cycle time depends on the acceleration, deceleration and maximum speed of the ropeway.

3.2 Aerial Ropeway

Aerial ropeways consist of two stations and free hanging gondolas which are attached to two types of wire ropes: a towing rope and bearing cables (SBN, 2007). Bearing cables are heavy ropes which are dimensioned to carry the load of the gondolas. These rigid ropes are mounted in the upper and lower station. The mounting of the bearing cables on several pylons along the route depends on the track length and the area that has to be crossed. The towing rope has a much smaller dimension; it has to forward the gondolas. The tensioning device is installed in the lower and the drive system is installed in the upper station of the ropeway. The number of gondolas depends on the maximum load of the ropeway. In addition to the possibilities of a pulsing or continuous movement of the towing rope, aerial ropeways are sub-divided into two different types: circulating and reversible ropeways.

Circulating aerial ropeways have a circumferential towing rope that underlies a constant movement. Thus, there are two synchronized tracks: up- and downstream. According to this, both the gondolas and the driving cable have to pass a disc at each end of the track to get on the opposite track. The diameter of this disc depends on the dimension of the gondolas. Regarding a 40 ft. marine container, that disc would have to be very large.

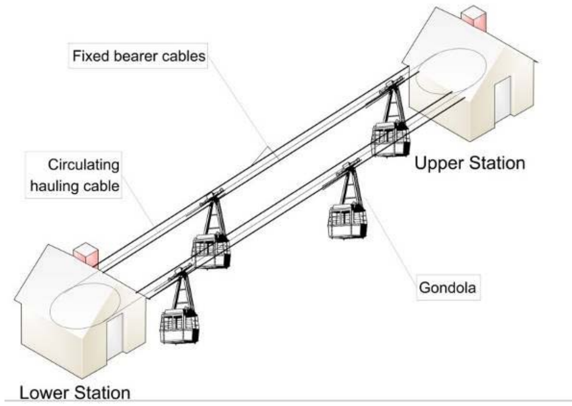


Figure 7. Circulating ropeway (Source: Authors)

One major benefit is the circulating aerial ropeway's possibility to transport several objects at the same time. Thus, this system offers a high output and therefore low cycle times but the payload of one gondola is usually low. The length and construction of the ropeway affects the number and maximum payload of the gondolas.

Reversible aerial ropeways function in a similar manner as the funicular ropeway. Instead of two tracks, there are installed two tracks of bearing ropes. Each of which is equipped with a gondola and a shared towing rope which is attached to each gondola or each of them is equipped with a towing rope. In the first case, one cable winch is installed within the upper station, which moves the towing rope. There is one gondola moving from the lower to the upper station and one gondola moving in the opposite direction at the same time. In the second case, each track has its own winch and can be moved independently. As a result, reversible ropeways are used to transport heavy single payloads over long distances.

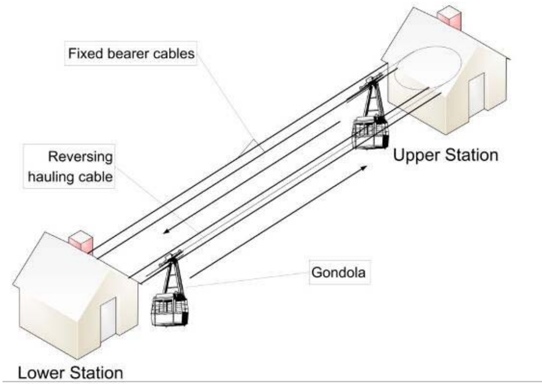


Figure 8. Reversible ropeway (Source: Authors)

This system can pick up high payloads and can overcome large barriers. Due to the fact that there is only one gondola on each track, the cycle time is longer than the cycle time of a circulating ropeway but similar to the cycle time of a funicular ropeway. To improve the cycle time regarding long distances, several ropeways might be installed and linked in a chain.

3.3 Conclusion regarding Ropeway for Marine Containers

The best technical system for an efficient transportation over short and medium distances may differ. Table 1 shows the results of a benefit analysis regarding the scenario mentioned in the introduction (short distance) compared with the demands, which have been posed in the second chapter.

For determination of the best solution for the regarded scenario the method of benefit analysis has been used. Therefore, ten weighting factors were defined on the basis of the functional and non-functional demands (cf. chapter 2). Following a definition of the importance of each factor, valued through a paired comparison, the target achievement of the four alternatives related to the weighting factors has been evaluated. The results of these ratings are shown in the white columns of Table 1. Consequently, these rankings have been multiplied with the ratings of the four technical alternatives as a measurement for the importance of a factor. The summation of each of the four alternatives is a measurement of the overall benefit of a system using the respectively technology. As illustrated in Table 1, the funicular ropeway fits best to the mentioned scenario. Regarding this scenario the high investment cost per meter of track is not as important as the cycle time of the system. In contrast to the demands of a short distance scenario, the best technology for a long medium scenario might not be the same because the investment costs per meter of track are much more important due to the distance the ropeway has to cover.

Table 1. Results of a benefit analysis (Source: Authors)

		Alternatives							
		Funicular Ropeway		Aerial ropeways				Multitrailer	
Weighting factor	Ranking			Circulating Ropeway		Reversible Ropeway			
	Scenario Eurogate	Rating	Value	Rating	Value	Rating	Value	Rating	Value
Investment costs (buffers and stations)	4	3	12	1	4	2	8	5	20
Investment costs (per meter of track)	2	1	2	2	4	3	6	5	10
Operating costs	7	4	28	4	28	4	28	1	7
Fully automatic transport	4	4	16	4	16	4	16	1	4
Transportation of 20/ 40ft. standard cont.	9	4	36	2	18	3	27	5	45
Maximum possible weight of load	2	5	10	1	2	3	6	4	8
Cycle time	9	4	36	5	45	3	27	1	9
Operational availability	7	4	28	4	28	4	28	5	35
Legal requirements	1	4	4	3	3	3	3	1	1
Possibility to overcome structural obstacles	6	2	12	4	24	4	24	1	6
∑			184		172		173		145
Legend: 1: very poor; 2: poor; 3: medium; 4: good; 5: very good									

In summary, a circulating aerial ropeway could be used to transport empty containers efficiently. Due to the large diameter of the disc, a large area has to be provided. As a result, the circulating ropeway is not an adequate solution for the transportation of marine containers. The reversible aerial ropeway has a better performance regarding payload and investment costs but it also runs slower. This technology is quite suitable to transport marine containers over a short distance but the best solution is the funicular ropeway. The funicular ropeway is an adequate solution for the transportation of marine containers. Its biggest disadvantage is the need of a bridge-structure but its relatively small investment costs and performance regarding heavy loads best fulfils the demands.

4. ROPEWAY FOR MARINE CONTAINERS

4.1 Funicular Ropeway for Marine Containers

The funicular ropeway can be built as described in Chapter 3.1 and shown in Figure 9.

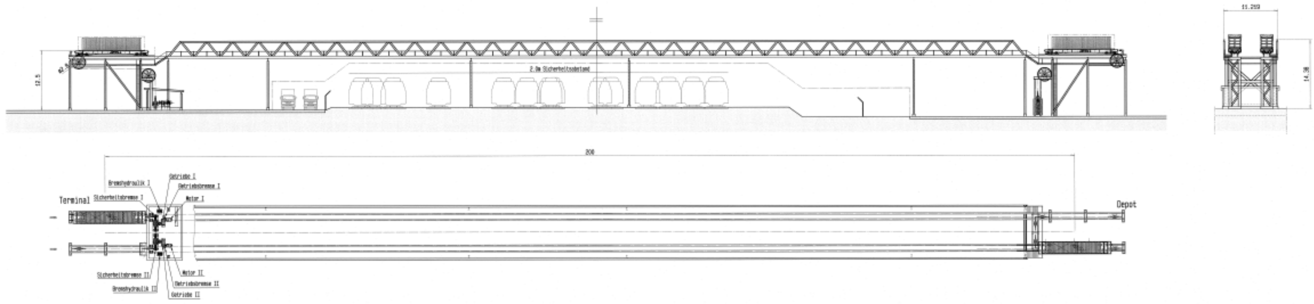


Figure 9. Funicular Ropeways (project partner)

In the following there will be an introductory description of the funicular ropeway for marine containers and a surrounding technical element: the buffer. The description of the ropeway refers to an offer of the world leading ropeway constructor.

4.1 Introductory description

Relating to the mentioned offer, two wagons that can be loaded optionally with either a 20ft. or a 40ft. container have to be installed. The maximum payload will be roughly 25 tons. One cycle, regarding one string and including loading and unloading as well as covering 200m, takes around 120 seconds. In this period approximate 60 seconds are intended for loading and unloading. The maximum speed of the ropeway will be about 5 meters per second. Thus, 30 wagons may be transferred per hour and sting, which enhances maximum transportation to 60 containers per hour. The calculated maximum average output is considered to be 750 tons per hour. To uncouple the cycle time of the ropeway system from the surrounding systems, a buffer has to be installed. The buffer will be described in the following chapter. To sum up, this offer meets or even exceeds all functional demands mentioned in Chapter 2. All non-functional demands have to be taken into account during final construction.

4.2 Buffer

The buffer has to be as large as possible to guarantee uncoupling of the systems and as small as possible to save space. Due to different space capacities regarding several seaports, the relation between storage capacity on the buffer and reduction of requirement of space may differ. Research has shown that the best size might have 3-4 separated buffer places. Thus, the buffer has to be separated into individually controlled segments to enable a steady supply of the ropeway. In this case, there has to be the same conveyor technology on the gondola and on the buffer to realise a safe and easy transfer between the buffer and the ropeway. The transfer between Reach Stacker / Van Carrier and ropeway has to be as simple as possible. The easiest way to supply both strings of the ropeway is to install one reception and one release station, for both strings together as shown in Figure 10.

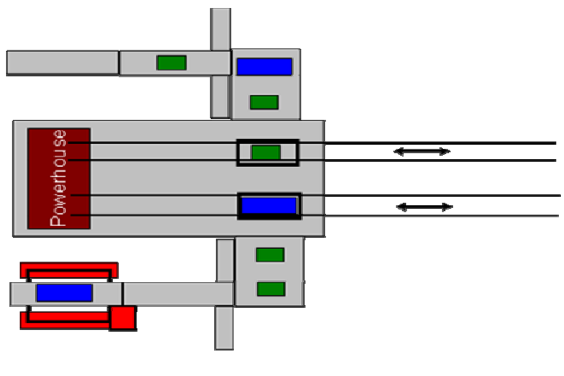


Figure 10. Big buffer solution (Source: Authors)

The cost of the buffer rises with each segment: cost and performance are interdependent. As shown in Figure 11., the minimalistic buffer version regarding the funicular ropeway for marine containers may consist of only one buffer place per string. This place is concurrent with the wagon. In this case, there would not be any conveying technology needed. If container A has been transferred, first a Van Carrier or Reach Stacker has to unload it from the wagon before container B, that has to be transferred, can be supplied. In each specific case, the user has to define buffer size and required performance to equip the ropeway for marine containers at best.

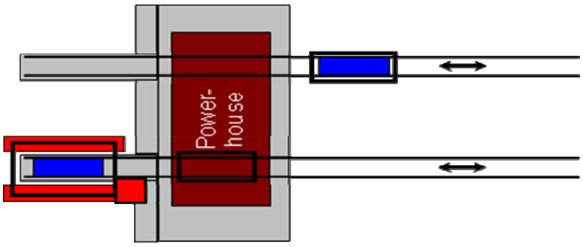


Figure 11. Small buffer solution (Source: Authors)

5. INFLUENCE ON EXISTING INFRASTRUCTURE, CONCLUSION AND OUTLOOK

The existing means of transportation, road and rail as well as the feeder traffic with ships are operating almost everywhere at their limit. There is a need for an alternative transportation module at many container terminals worldwide. With an air-bound transportation module, the crossing of land and water is possible for both short and long distances.

The ropeway for marine containers can transport 20 and 40 ft. containers with a weight of up to 20 tons fully automatically over a distance of 200 meters. This concept mainly bases on special requirements regarding the projects' scenario. But having included also general aspects like the possible adoption of other scenarios, the presented results may be transferred to other scenarios. Further possible applications for the ropeway for marine containers are container-terminals with strong hinterland connections (Europe, the USA) as well as container-terminals with strong internal traffic (e.g. northeast Asia). A second phase of the project could be initiated to realize the ropeway for marine containers. The final definition of the ropeway, especially the gondola or wagon, and the buffers has to be done.

5.1. Influence on existing terminal infrastructure

The ropeway for marine containers is an enormous technical challenge. However, the technology regarding e.g. the automatic transfer from the buffer to the gondola or wagon must be developed; but this technology can later be transferred to other means of transport, e.g. straddle carriers and gantry cranes. In addition, the transport of the containers via ropeway has to be integrated within existing material and information flows. For optimal utilization of the ropeway buffers must be configured with regard to the special requirements. The internal terminal-software that steers all movements on the terminal has to be suited to the new infrastructure module. To simplify matters, an assigned storage place on the opposite area might be equipped with a sub-movement that includes the transportation via ropeway. The design depends on the needs of the user.

In terms of the mentioned project, the container-terminal will replace the existing trailer traffic through a ropeway for marine container. Therefore, the existing material flow between the terminal and the outside storage need not be changed but the steering has to be suited.

5.2. Influence on general public

The employment of a ropeway for marine containers has extensive effects on the infrastructure close to container terminals. Hence the existing infrastructure at a seaport would be less used: Supplying the ropeway with containers inhibits their transportation via truck. Regarding other scenarios like the possible hinterland-connection via ropeway, this innovative infrastructure module also effects the planning of road and bridge networks in the region. Also in regard to energy efficiency and noise pollution the ropeway for marine containers offers advantages. The investigations have shown that operationally the investment can reduce fixed running costs. This infrastructure module could be used to improve the national goods traffic network in the future.

5.3. State of affairs

In the second phase of the project, the ropeway for containers will be designed to be installed at the locality as described in Chapter 1. The first step of the project was a feasibility study to see if a solution already existed and if a ropeway for marine containers could work efficiently. The study analysed the circumstances, investigated existing solutions, defined the demands on the technical system, found a potential provider, discussed the solutions with the partners and calculated the economic efficiencies of the ropeway for marine containers. To follow up, the ropeway for marine containers has to be built and installed in the second project phase.

Afterwards, the ropeway for containers has to be transferred to other scenarios like the sea port-hinterland connection or a scenario where it has to transport containers above different surfaces like water.

To sum up, the investigation showed that it is possible to use a ropeway for the transportation of containers. This technology guarantees a fully automatic connection of separated areas. Additionally, a ropeway may be used to transfer containers over medium or long distances as for example in order to connect a seaport and the hinterland. In comparison to today's technologies like trucks or special freight vehicles, the ropeway emits less CO_2 and less noise. Future tasks are the final construction of the wagon and the implementation of the system in Hamburg. Further on, new scenarios have to be defined and investigated to support the diffusion of this innovative infrastructure module.

6. REFERENCES

Gardiner, N. (2009), Container Market 2009/10, Annual Review and Forecast - Incorporating the Container Forecaster - 3Q09, Drewry

Maaß, S., Kopp, M. (2010). Erste Anzeichen für Erholung im Hafen, *DIE WELT* submitted on 23th January 2010

Schweizerischer Bundesrat (2007). Verordnung über Seilbahnen zur Personenbeförderung. [WWW] <URL: http://www.seilbahn.net/snn/konfig/uploads/doku/103.pdf> [Accessed 26 February 2010.]

SBN: Seilbahn.Net Information platform (2007). Operations manager conference. [WWW] <URL: http://www.seilbahn.net/snn/konfig/uploads/doku/96.pdf> [Accessed 26 February 2010.]

Port of Hamburg (2010): Annual Port of Hamburg press conference. Hamburg, submitted on 5th February 2010.

AN ANALYSIS OF THE COMPETITION OF PORTS IN SHANGHAI INTERNATIONAL SHIPPING HUB

Sufen Chen* , Louie Nan Liu, and Guoqing Liu

School of Management
Zijingang Campus，Zhejiang University
Hangzhou 310058 China
E-mail: chen.sufen@163.com

Abstract: Globalization of world economy has led to a rapid growth of container transportation as well as container ports. In China, Shanghai and Ningbo ports have experienced an explosive growth rate based on their strong economic hinterlands. The objective of this research is to analyze the spatial structure of the ports cluster and the market share of the ports at their conjunct hinterlands qualitatively and quantitatively with a case study of Shanghai International Shipping Hub with its two main ports—Shanghai and Ningbo ports over the period between 1995 and 2004. Our findings suggest that Ningbo port has now emerged as a serious contender to Shanghai port in Shanghai International Shipping Hub but its position has not been a dominant one. It shows the spatial structure of Shanghai International Shipping Hub and forecasts the future structure of double-hub port mode, which has profound implications for China's water transport restructuring.

1. INTRODUCTION

Along with the rapid development of China's economy, both domestic trade and foreign trade have been growing enormously. And freight transport with container becomes increasingly popular because it is not only convenient and safe but also fast. Over 90% of foreign trade logistics volume depends on ports in China and container transport acts as the major form. Since 2002, China's foreign trade development has been keeping a 20% growth. The foreign trade surmounted US$1 trillion in 2004. Moreover, 27% of the cargoes come from the region of Yangtze River Delta. These containers in this region are mainly transported through Shanghai port and Ningbo port of Shanghai International Shipping Hub. The massive export surge in China over the last decade has driven the expansion of container ports. Therefore, it is of significantly important to study on the competitiveness of ports.

Shanghai International Shipping Hub has been officially started from January 1996, consisting of the coastal ports of Jiangsu and Zhejiang Provinces with the key port at Shanghai Municipality (Figure 1). The port cluster of Shanghai International Shipping Hub has not been well defined up to now. According to Shanghai ports cluster stated in the previous years, Shanghai International Shipping Hub here is defined to include most coastal ports of Yangtze River Delta, which are Shanghai port, Ningbo port, Suzhou port including Taicang, Changshu and Zhangjiagang ports, Yangzhou port, Jiangyin port, Zhenjiang port, Nanjing port, Nantong port, Tàizhou port, Zhoushan port, Wenzhou port and Tāizhou port, totally twelve ports.

Shanghai port handled 14.55 million TEUs in 2004, with an annual average increase of 28.1% since 1990, ranking number one in China and number three in the world respectively. Ningbo port is one of China's deep-water ports, enjoying unique natural conditions. Since 1990s, Ningbo port has grown with an amazing surge of cargoes. In 2004 the TEU's handling reached 4 million, two hundred times of the year 1990, with an annual average increase of 45.6%. Shanghai port and Ningbo port compete fiercely for scrambling the cargoes at the hinterlands, especially the conjunct hinterlands of Zhejiang province. With the continuing development of intermodal transport of the container in China, the traditional concept of conjunct hinterland becomes to vanish. Shanghai port and Ningbo port have more and more conjunct hinterlands. According to Hoare (1986), the competition between Shanghai port and Ningbo port will be increasingly fierce.

In China, in response to the significant growth of container ports, there have been some research studies on the spatial structure of container ports. Cao et al. (2004) indicate the spatial structure of China's container port system and the competitive pattern. The Chinese Government has also been restructuring the China's transport space in order to reflect the growth of container ports brought about by world economic globalization. General Layout and Plan of Coastal Port of China have begun to be drawn up by Minister of Transport and National Committee of Development and Reform since 1993. Yet the methods of all these studies tend to be qualitative and descriptive in nature.

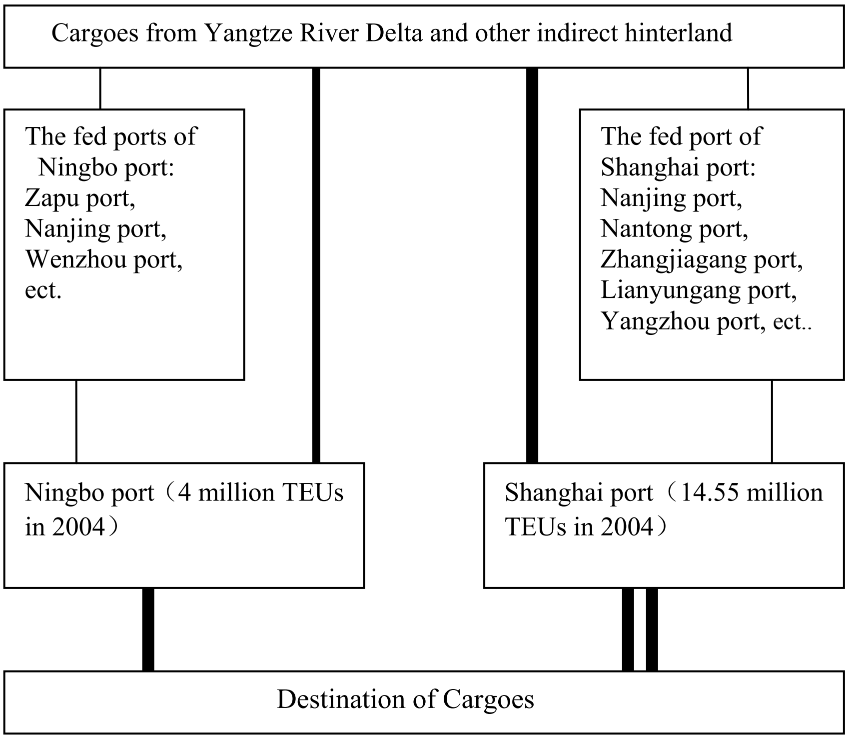


Figure 1. The Present Container Port System of Shanghai International Shipping Hub

Internationally, there was few research found on the container port system in the early stage. Taaffe et al (1963) demonstrates that a centralization trend of container ports during their growth results in the minor ports' peripherization even disappearance. Barke (1986) and Slack (1990) adds that decentralization phenomenon comes after the process of centralization of the ports. Hayut (1981) brings forward the stage development model of the container port system, which is most authoritative research on the container port system. Hayuth considers the development of container port system experiences preparatory stage, container-used stage, ports centralization stage, major and center stage and challenge of the surrounding minor ports stage. In the first four stages, the container port system trends to centralize until the container major port comes forth. In the fifth stage, diseconomies of scale effect will appear with the further expansion of the major port. Thus the dominant situation of the major port faces the challenge from a scatter of minor or new-rising ports.

In the research of Airriess (1989), Kuby, Reid (1992), Hoyle and Charlier (1995), the process of centralization does take place obviously. And some other researchers challenge the assumption of Hayut and makes amendments. Notteboom (1997) indicates containerization does not always result in the process of ports centralization through the research on the development of the European container system. Starr (1994) provides the solution, which Hayuth Model cannot answer that two major ports may coexist in one region. O'Mahony (1998) finds there exists large deep direct - sea container port between the major ports and the fed ports.

Therefore, this study on the competitiveness of ports with a case study of Shanghai International Shipping Hub takes quite an important significance. This paper focuses on the methods to measure the degree of concentration of a container port system, and analyze the market share of the ports at their hinterlands. This paper studies the competitiveness of the two main ports (Shanghai port and Ningbo port) in Shanghai International Shipping Hub over the period between 1995 and 2004. First, the port's performances in container volume are usually considered as the sum of its competitive positions. Through Herfindahl-Hirschman Index (HHI) model, the paper studies the port's position in a ports cluster. Second, the gains or losses of market share by individual port are often considered as an index of their competitiveness. The shift-share analysis breaks down the change of container volume in a port into two components: shift effect and share effect.

The organization of this paper is as follows. Following the introduction, Section two presents the HHI models for analyzing quantitatively and qualitatively the spatial structure of the port cluster, and shift-share analysis and location quotient for analyzing the increasing container volume of the ports. In section three a case study of Shanghai

International Shipping Center is applied. It presents the results of the application of HHI model, and shift-share analysis and location quotient. At last, section four gives some conclusions.

2. RESEARCH MODELS

2.1 Methods to Measure the Degree of Concentration of a Container Ports System

Herfindahl-Hirschman Index – HHI

HHI stands for Herfindahl-Hirschman Index, a commonly accepted measure of market concentration. The HHI is often used to characterize the distribution of a variable of interest by measuring its degree of concentration across units, such as firms, teams or households (Owen et al., 2005). It is calculated by squaring the market share of each firm competing in a market, and then summing the resulting numbers. The HHI number can range from close to zero to 1. The HHI is expressed as:

$$HHI = s_1^2 + s_2^2 + s_3^2 + ... + s_n^2 \tag{1}$$

Where s_i is the market share of the i^{th} firm in an industry comprising n firms. Given that $0 \leqslant s_i \leqslant 1$, for all i, $s_1 + s_2 + s_3 + ... + s_n = 1$ and then 1/n≤HHI≤1. The lower bound, 1/n, corresponds to a situation of equal market shares for each of the n firms, whereas the upper bound of unity is the case of pure monopoly. In the context of measuring the competitiveness of a port in a port cluster, the HHI can be applied to port's shares of container throughputs over a port cluster, and the interpretation of the upper and lower bounds carries over in this case. In a 'perfectly balanced' port cluster with n ports, each port would averagely share the total container throughputs. At the other extreme, if one port owns all the container throughputs of a port cluster, then HHI = 1, as in the case of the monopolist's market share. Hence, equation (1) becomes as follows:

$$D = \frac{\sum_{i=1}^{n} TEU_i^2}{\left(\sum_{i=1}^{n} TEU_i\right)^2}, \text{ of which } \frac{1}{n} < D < 1 \tag{2}$$

Where D represents the degree of centralization of the system of the container ports and TEU_i is the container throughputs of port i, and n equals the number of the ports in the ports cluster.

Generally, there exists concentration in a port cluster when HHI is larger than 0.1; and it indicates that a port cluster is highly centralized when HHI is larger than 0.18.

Shift-Share Analysis

Shift-share analysis has been used extensively to analyze the differences between regional and national growth rate in variables such as export growth, employment and productivity (Wilson et al., 2005) . In this study, we focus on container throughput growth over a period of time where the 'regions' are the ports (Shanghai port, Ningbo port, Suzhou port , Yangzhou port, Jiangyin port, Zhenjiang port, Nanjing port, Nantong port, Tàizhou port, Zhoushan port, Wenzhou port and Tāizhou port) of Shanghai International Shipping Hub, and the 'nation' is the combined group of these ports. In this context the objective is mainly to compare the Shanghai port's performance in terms of container throughput volume against the reference port cluster which includes its competitor Ningbo port.

The gains or losses of market share by individual port are often considered as an index of their competitiveness. The shift-share analysis divides the change of container throughputs in a port into two components: shift effect and share effect. The growth rate of the reference ports cluster is considered to be the share effect. The difference between a port's performance and the part of the total change is referred to as the shift effect. The shift effect or share effect in this paper is measured in TEU of container throughputs. A net positive shift implies improvements in competitiveness relative to the ports cluster as a whole and a negative value constitutes deterioration in competitiveness. Thus, the formula is presented as follows:

$$ABSGR_i = TEU_{it_1} - TEU_{it_0} = SHARE_i + SHIFT_i \tag{3}$$

$$SHARE_i = \left(\frac{\sum_{i=1}^{n} TEU_{it_1}}{\sum_{i=1}^{n} TEU_{it_0}} - 1 \right) * TEU_{it_0} \tag{4}$$

$$SHIFT_i = ABSGR_i - SHARE_i \tag{5}$$

Where $ABSGR_i$ represents the absolute growth of container throughput of port i over the period (t_0, t_1). $SHARE_i$ is the share increase of container throughput of port i during the period (t_0, t_1). TEU_{it_1} and TEU_{it_0} are the container throughputs of port i in the period t_1 and t_0, respectively, and n again, is denoted as the number of the ports in the port cluster.

Moreover, we define here $VOLSHIFT$ as the shift between the ports. With $VOLSHIFT$, we can get how competitive between the ports. The bigger value of $VOLSHIFT$ is, the fiercer competitiveness between the ports exists.

$$VOLSHIFT = \frac{\sum_{i=1}^{n} \left| SHIFT_i \right| - \left| \sum_{i=1}^{n} SHIFT_i \right|}{2} \tag{6}$$

Where: $VOLSHIFT$ = shift between ports in a port cluster

n = the number of ports in a port cluster

2.2 Quantitative Analysis of the Market Share in Terms of Different Regions

Location Quotient of Port Hinterlands Model

Location Quotient (*LQ*) is a ratio measure of the concentration of a cluster in a particular location relative to the national average. So, *LQ* is a measure of an industry's level of concentration within a location. The industry's location quotient for province i, is the ratio of the industry's share of total province employment relative to its share of total national employment. Thus, the formula for the location quotient in a province (can be uniformly applied to economic area or metropolitan area) is as follows:

$$LQ_(= \frac{E_{i,j} / E_i}{E_{n,j} / E_n} \tag{7}$$

Where $E_{i,j}$ refers to province i 's employment in industry j,

E_i is the total employment in province i,

$E_{n,j}$ is total nation's employment in industry j ,

E_n is the total employment in the country.

A location quotient more than one indicates that the industry enjoys a greater share of employment in the province than it does nationally, while a location quotient less than one indicates the opposite.

The selection of the base or distribution used in location quotients is subject to choice (Barber, 1988). In this study, it makes sense to compare the concentration of various regions (hinterlands) to the total container throughput in all regions (hinterlands). In this instance, location quotient method is used in this research to analyze the hinterlands connected to the neighboring ports. The model is established as below:

$$LP_{ij} = \frac{h_{ij} \Big/ \sum_i h_{ij}}{\sum_j h_{ij} \Big/ \sum_i \sum_j h_{ij}} \tag{8}$$

Where: LP_{ij} = the relative degree of concentration of the cargoes of region j to port i

h_{ij} = the cargoes of region j transferred through port i

$\sum_i h_{ij}$ = the total cargoes of region j

$\sum_j h_{ij}$ = the cargoes of all the regions transferred through port i

$\sum_i \sum_j h_{ij}$ = the total cargoes of all the regions

The result LP_{ij} greater than 1 indicates that the cargoes of region j are mainly transferred through port i, this is to say region j is a hinterland of port i . The greater figure of LP_{ij} shows that port i has stronger competitiveness in region j .

3. CASE STUDY

3.1 Degree of Concentration of Container Port Cluster of Shanghai International Shipping Hub

Results of HHI Method

This research collects the annals from 1995 to 2004 of the 12 ports of Shanghai International Shipping Hub, and uses the HHI method to analyze the degree of concentration of Shanghai port, Ningbo port and both the two ports. The results are presented in Table 1.

Table 1 Analysis of Concentration of Shanghai International Shipping Hub

	1995	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*	*2002*	*2003*	*2004*
HHI	0.556	0.595	0.625	0.639	0.611	0.607	0.579	0.573	0.554	0.545
ROSH	73.4%	76.3%	78.4%	79.2%	77.1%	76.8%	74.6%	73.9%	72.2%	71.1%
RONB	7.7%	7.8%	7.9%	9.1%	11.5%	12.3%	14.3%	15.9%	17.7%	19.6%
ROTP	81.2%	84.1%	86.3%	88.3%	88.7%	89.1%	88.9%	89.8%	89.9%	90.7%

Source: (1) Yearbook of Port, China (1999-2004); (2) Annals of the Ports.

ROSH (Ratio of Shanghai) stands for the ratio of the container throughputs of the ten years of Shanghai port out of the total container throughputs of port cluster of Shanghai International Shipping Hub, and hereafter called Ratio of Shanghai; RONB (Ratio of Ningbo) stands for the ration of the container throughputs of the ten years of Ningbo port out of the total container throughputs of the port cluster of Shanghai International Shipping Hub, and hereafter called Ratio of Ningbo. ROTP（Ratio of Two Ports）means the ratio of the sum of the container throughputs of the two ports(Shanghai port and Ningbo port) out of the total container throughputs of port cluster of Shanghai International Shipping Hub, and hereafter called Ratio of the ports.

Table 1 shows the features of statistics. The percentage of container throughput by port is contained in Line 2, 3 and 4. These values are used to derive the HHI for the shipping Hub. They confirm that in Shanghai International Shipping Hub the container volume is concentrated in Shanghai port: 71.1% of the container volume is handled over in Shanghai port in 2004. Furthermore, the degree of concentration of the ports cluster tends to rise firstly between 1995 and 1999 and then declines between 2000 and 2004. The decentralization is the result of the growth differences between Shanghai port (-2.3% in ten years) and Ningbo port (11.9%).

From Table 1, HHI of Shanghai International Shipping Hub is between 0.545 and 0.639, greater than 0.18, indicating that the shipping Hub is in the stage of high degree of concentration. At the same time, HHI first rises from 0.556 (in the year of 1995) to 0.639 (in 1999) then falls to 0.545 (in 2004). The port cluster experiences the process of concentration and decentralization. The reason is that the dominant port, Shanghai port, does not always keep the annual growth rate which is greater than the average annual growth rate of the ports cluster. For example, from 2000 to 2004, Shanghai port loses its 1154.0 thousand TEU (Table 2); Ningbo port, Yangzhou port, Tàizhou port keep a growth rate which is greater than the average growth rate, especially Ningbo port, during 2000 to 2004, gains the shifts of 1480.2 thousand TEU, accounting for 94.4% of the net shifts of the ports cluster.

Therefore, the speedy growth of container throughputs of Ningbo port is the main factor that leads to the decrease of HHI of the ports cluster from 1998 (table 1). It is expected that with further growth of container throughputs of Ningbo port, the figure of HHI index of the port cluster will further drop. The ports cluster tends to decentralize.

The results of Shift-share Analysis

The results of shift-share analysis for Shanghai International Shipping Hub for 1995-2004 are illustrated in Table 2. The subsequent paragraphs describe each of the three shift-share components in more detail.

In both periods, all the ports of Shanghai International Shipping Hub have a net growth. Shanghai port and Ningbo port are the best two performers. Table 2 provides a representation of the evolution over the two periods of the net shifts for the ports' growth. To facilitate interpretation, the data are left in absolute terms (ten thousand TEU).

Firstly, the profiles in Table 2 suggest that Ningbo port has now emerged as a serious competitor with an increasing number of sustained positive net shifts in container throughput from 1995 up to 2004, but its position has not been a dominant one in Shanghai International Shipping Hub. Secondly, in the second period, Shanghai port recorded the largest negative net shifts (-115.40), however it remains its dominant position.

Table 2 Analysis of Shift-share of Container Port Cluster of Shanghai International Shipping Hub
(unit: ten thousand TEU)

	1995～1999			*2000～2004*		
	Net growth	*Shares*	*Shifts*	*Net growth*	*Shares*	*Shifts*
Shanghai port	249.90	230.73	19.17	894.25	1009.60	-115.40
Ningbo port	44.14	24.18	19.96	310.33	162.01	148.02
Nanjing port	0.78	22.53	-21.75	28.40	36.52	-8.12
Nantong port	7.08	13.18	-6.10	10.54	32.74	-22.20
Zhenjiang port	1.90	2.88	-0.99	10.58	10.29	0.29
Suzhou port	1.75	16.33	-14.58	28.03	35.44	-7.41
Zhoushan port	0.08	0.00	0.08	2.98	3.54	-0.56
Yangzhou port	0.98	0.00	0.99	10.99	3.80	7.18
Wenzhou port	2.68	2.92	-0.23	13.90	13.33	0.56
Jiangyin port	0.73	0.00	0.73	1.73	2.50	-0.77
Tàizhou port	3.63	0.00	3.63	1.67	0.66	1.02
Tāizhou port	0.43	1.30	-0.86	1.90	4.22	-2.32
Total	314.08	314.08	0.00	1315.25	1315.25	0.00
Net shift			44.50			156.77

Source: (1) Yearbook of Port, China (1999-2004). (2) Annals of the Ports.

Market Share Analysis of Zhejiang Province

Location quotient analysis of the market share of Ningbo port and Shanghai port

The research uses location quotient method with the annual report of cargoes of the regions of Zhejiang province from 2001 to the first half year of 2005, to analyze the competitiveness between Shanghai port and Ningbo port in Zhejiang province.

For Shanghai port, the location quotients reveal a concentration in Hangzhou, Jiaxing, Huzhou, Jinhua, and Quzhou. However, for Ningbo port, the distribution is concentrated in Ningbo, Tāizhou and Zhoushan. Table 3 indicates Hangzhou, Jiaxing, Huzhou, Jinhua and Quzhou are the direct hinterlands of Shanghai port, and their cargoes transferred through Shanghai port. Ningbo, Tāizhou and the region of Zhoushan are the direct hinterland of Ningbo port in Zhejiang province. The cargoes of the region of Shaoxing in 2001 are transferred outside of Shanghai port and Ningbo port, and in the following years the region of Shaoxing becomes the direct hinterland of both ports, showing the ports compete strongly in the region of Shaoxing. The regions of Wenzhou and Lishui are not the direct hinterland of Shanghai port and Ningbo port.

Table 3 Result of Location Quotient of Cargoes of the Regions of Zhejiang Province (2001-the first half year of 2005)

	Shanghai Port-2001			Shanghai port-2002			Shanghai port-2003			Shanghai port-2004			Shanghai port-2005		
	Imp	*Exp*	*Ave*	*Imp*	*Exp*	*Ave*	*Imp*	*Exp*	*Ave*	*Imp*	*Exp*	*Ave*	*Imp*	*Exp*	*Ave*
Hangzhou	1.46	1.39	1.43	1.45	1.44	1.45	1.30	1.46	1.38	1.28	1.44	1.36	1.35	1.44	1.39
Wenzhou	0.20	0.70	0.45	0.26	0.71	0.49	0.36	0.78	0.57	0.48	0.74	0.61	0.43	0.72	0.57
Jiaxing	2.11	1.55	1.83	2.34	1.62	1.98	2.77	1.72	2.25	2.81	1.84	2.32	3.14	1.92	2.53
Huzhou	2.34	1.52	1.93	2.30	1.62	1.96	1.90	1.69	1.80	1.89	1.77	1.83	1.99	1.85	1.92
Shaoxing	0.62	1.13	0.88	0.96	1.11	1.03	0.95	1.16	1.06	1.17	1.18	1.17	0.93	1.22	1.08
Jinhua	1.01	1.21	1.11	1.15	1.54	1.35	1.33	1.03	1.18	0.85	1.06	0.96	1.59	1.03	1.31
Quzhou	0.98	1.10	1.04	1.23	1.10	1.16	0.90	1.17	1.03	0.94	1.26	1.10	1.31	1.19	1.25
Zhoushan	0.31	0.43	0.37	0.12	0.33	0.23	0.82	0.28	0.55	0.60	0.27	0.43	0.83	0.38	0.61
Lishui	0.20	0.86	0.53	2.32	0.82	1.57	2.93	0.77	1.85	0.95	0.73	0.84	0.42	0.67	0.55
Tāizhou	0.36	0.50	0.43	0.56	0.50	0.53	0.32	0.47	0.40	0.28	0.44	0.36	0.34	0.38	0.36
Ningbo	0.19	0.46	0.33	0.27	0.48	0.37	0.25	0.46	0.36	0.29	0.44	0.36	0.30	0.42	0.36
	Ningbo port-2001			Ningbo port-2002			Ningbo port-2003			Ningbo port-2004			Ningbo port-2005		
Hangzhou	0.35	0.34	0.34	0.36	0.33	0.35	0.38	0.34	0.36	0.45	0.30	0.37	0.47	0.30	0.38
Wenzhou	0.09	0.87	0.48	0.15	1.02	0.59	0.18	1.08	0.63	0.18	1.20	0.69	0.11	1.22	0.67
Jiaxing	0.09	0.05	0.07	0.05	0.07	0.06	0.07	0.08	0.07	0.11	0.07	0.09	0.06	0.07	0.06
Huzhou	0.00	0.16	0.08	0.03	0.11	0.07	0.00	0.13	0.07	0.00	0.11	0.05	0.00	0.09	0.04
Shaoxing	0.83	0.76	0.80	0.96	0.86	0.91	1.29	0.83	1.07	1.28	0.84	1.06	1.37	0.81	1.09
Jinhua	0.51	0.67	0.59	0.42	0.75	0.59	0.44	0.91	0.68	0.31	0.92	0.62	0.55	0.90	0.73
Quzhou	0.00	0.33	0.16	0.00	0.59	0.29	0.12	0.57	0.35	0.40	0.66	0.53	0.54	0.79	0.66
Zhoushan	0.75	1.59	1.17	0.32	1.58	0.95	0.60	1.65	1.12	0.18	1.56	0.87	0.26	1.29	0.77
Lishui	0.00	1.39	0.65	0.15	1.39	0.77	0.00	1.31	0.65	0.12	1.37	0.75	0.11	1.41	0.76
Tāizhou	0.53	1.94	1.24	0.65	1.83	1.24	0.64	1.78	1.21	0.59	1.80	1.20	0.54	1.85	1.19
Ningbo	2.82	2.13	2.47	2.29	1.92	2.11	2.12	1.80	1.96	2.02	1.82	1.92	1.87	1.79	1.83

4. CONCLUSION

This paper is based on much practical investigation by the authors in Shanghai port and Ningbo port. It has described the present situation of Shanghai International Shipping Hub with its two main ports—Shanghai port and Ningbo port, and employed the HHI model and shift-share technique to analyze quantitatively the degree of concentration of port system and the market share of port at the hinterlands.

The methodologies are applied to a case study of Shanghai International Shipping Hub. The container volume of the ports cluster is concentrated in Shanghai port. Shanghai port remains its dominant position. It has also examined Shanghai port and Ningbo port's performance in growth of container volume between 1995 and 2004. Our findings suggest that Ningbo port has now emerged as a serious contender to Shanghai port in Shanghai International Shipping Hub but its position has not been a dominant one. Ningbo port has been keeping an increasing growth with a sustained positive net shift between 1995 and 2004. Ningbo port's competitiveness is expected to be sustainable in the future. Its perfect natural superior conditions together with strongly developed hinterlands put Ningbo port in a favorable position. It is expected to become another hub of Shanghai International Shipping Hub. Moreover, Shanghai port is facing increasing competition from its competitor, Ningbo port, in the container volume growth. This is reflected in the significant positive net shifts of Ningbo port.

The future Double Hub Port Mode of Shanghai International Shipping Hub is put forward in this paper based on the study of Starr (1994), as shown by Figure 2. Shanghai Port will be reinforced by completion of the construction of deep-water Yangshan Port. Together with the advantages of the unique hinterlands and the huge container volume, Shanghai will certainly become the most important intersection in Asia, and also it will be possible to transship Korea's and Japan's port cargoes via Shanghai Port in the future. Shanghai will surely become the international hub in northeast Asia just as Hong Kong port and Singapore port.

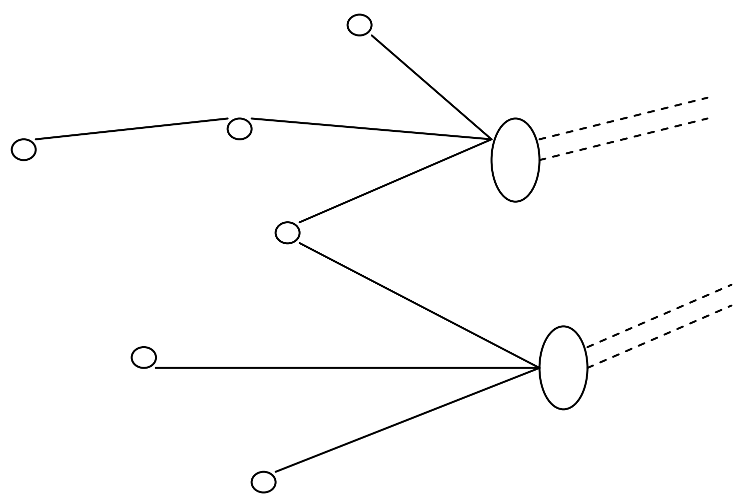


Figure 2. Double Hub Port Container Port System

The findings in this paper have important policy implications for China's port development. Hangzhou Bay Bridge has been officially started to construct since June 2003. The bridge connects with the Haiyan County of Jiaxing region with the northern end and Cixi City of Ningbo with the southern end. With Hangzhou Bay Bridge open to traffic, Ningbo will links with Shanghai, southern part of Jiangsu province, Taizhou, Wenzhou and even the southern cities of Fujian province. This will deeply impact the development of hinterland for Ningbo. Ningbo will expand its hinterland to Shanghai, northern cities of Zhejiang province and southern cities of Jiangsu province. On the other hand, Shanghai port will also extend its hinterland to the cities of Zhejiang province. Thus, Hangzhou Bay Bridge will have great impact on the competition between Shanghai port and Ningbo port. Additional research could be done in these areas.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funding supports of the Zhejiang Provincial Social Science Foundation (07HZC312Z) and the New Century Excellent Talent Support Program (NCET05-0529) of the Ministry of Education of China.

5. REFERENCES

Airriess C.A.(1989). The Spatial Spread of Container Transport in a Developing Regional Economy. *Tansportation Research Part A*, 23: 453-461.

Cao Y.H., Li J.H. and Chen W. (2004). The Spatial Structure and the Competition Pattern of the Container Port System of China. *Acta Geographica Sinica*, 59: 1020-1027.

Hayut Y. (1981). Containerization and the load center concept . *Economic Geography*, 57(2):160-176.

Hoare, A.G. (1986). British Ports and Their Export Hinterland: a Rapidly Changing Geography. *Geografiska Annaler*, 688B: 29-40.

Hoyle B.S. and Charlier J.J. (1995). Inter-port Competition in Developing Countries: an East African Case Study. *Journal of Transport Geography,* 3(2): 87-103.

Kuby M. and Reid N. (1992). Technological Change and the Concentration of U.S. General Cargo Port System [J]: 1970-88. *Economic Geography*, 68(3): 272-289.

Notteboom T.E. (1997). Concentration and Load Center Development in the European Container Port System[J]. *Journal of Transport Geography,* 5(2): 99-115.

Slack B. (1990). Intermodal Transportation in North America and the Development of Inland Load Centers[J]. *Professional Geographer*, 42(1):72-83.

Starr J.T. (1994). The Mid-Atlantic Load Center. *Maritime Policy and Management*, 23(3): 219-227.

Taaffe Edward J., Morrill R.L. and Gould P.R. (1963).Transport Expansion in Underdeveloped Countries:a Comparative Analysis. *Geographical Review*, 53: 503-529.

P. Dorian Owen, Michael Ryan and Clayton R. Weatherston (2005).*Measuring Competitive Balance in Professional Team Sports Using the Herfindahl-Hirschman Index: A Cautionary Note*. http://www.business.otago.ac.nz/econ/personal/do_files/Competitive%20Balance_NZAE%20(3).pdf. Accessed July 8, 2010.

Peter Wilson, Ting Su Chern, Tu Su Ping and Edward Robinson. *A Dynamic Shift-Share Analysis of the Electronics Export Market 1988-2001: Can the NIEs Compete with China?* SCAPE Working Paper Series, Paper No. 2005/07, May 2005. http://nt2.fas.nus.edu.sg/ecs/cent/Working%20Papers.htm. Accessed July 8, 2010.

Barke M. (1986). *Transport and Trade.* Oliver & Boyd, Edinburgh.

Gujarati D.N. (1995). *Basic Econometrics, 4th Edition*. The McGraw-Hill Compannies.

Li Z.N. (1992). *Econometric: Method and Application*. Press of Qinghua Univerity, Beijing.

Michael D. Intriligator (1997). *Econometric Models,Techniques and Applications (2nd Edition)*，Prentice-Hall Inc.

O' Mahony H. (1998). *Opportunities for Container Ports: a Cargo System Report.* London, IIR Publications.

Robert S. Pindyck and Daniel L. Rubinfeld (1999). *Econometric Models and Economic Forecast*. China Machine Press, Beijing.

Sargent A.J. (1938). *Seaports and Hinterlands.* London:Black.

Taaffe Edward J and Howard L Gauthier (1996). *Geography of Transportation (2nd Edition).* Prentice Hall Inc.

USING DEA TO MEASURE THE SEAPORTS EFFICIENCY OF CHINA AND FIVE MEMBER COUNTRIES FROM ASEAN

Mu Mu Han, Lin Guolong, Yang Bin

Logistics Research Center, Shanghai Maritime University, Shanghai 200135, China
mumuhan@gmail.com

Abstract: In this paper, the efficiency and performance is evaluated for ports of China and five member countries(Singapore, Philippines, Brunei-Darussalam, Indonesia, Myanmar) from Association of Southeast Asian Nations (ASEAN).The aim of our study is to compare seaports situated on the maritime trade road between China and five member countries from ASEAN. Data was collected for 9 years (1999-2007) and a non-parametric linear programming method, DEA (Data Envelopment Analysis) is applied. This method is used to measure the efficiencies of ports in different countries, is constructed with three input factors (numbers of cranes, numbers of berths, quay length) and two output factors (Throughput, vessel calls).The goal of our study to estimate the performance level of the ports under considerations. This paper discusses the efficiency analysis, and gives specific direction for the inefficient ports to improve their operation efficiency possibly. This will help in proposing solutions for better performance and developing plans.

Key word -China and five number countries form ASEAN Seaports; Data Envelopment Analysis (DEA); evaluation; Seaports Efficiency.

1. INTRODUCTION

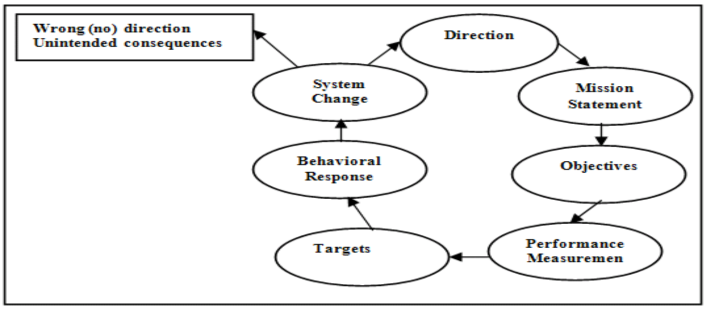
The globalization of the world economy has led to an increasingly important role for transportation. In particular, container transportation plays a key role in the process, largely because of the numerous technical and economic advantages it possesses over traditional methods of transportation. Standing at the crucial interface of sea and inland transportation, the significance of the container port and its production capabilities cannot be ignored. Compared with traditional port operations, containerization has greatly improved port production performance because of two reasons. To reap economies of scale and of scope, liner shipping companies and container ports are respectively willing to deploy dedicated container ships and efficient container handling systems. In so doing, port productivity has been greatly enhanced. On the other hand, many container ports no longer enjoy the freedom yielded by a monopoly over the handling of cargoes from within their hinterland. They are not only concerned with inter-port competition, under the orthodox microeconomic framework, is believed to provide an incentive to improve port performance. Productive efficiency, therefore, is a survival condition in a competitive environment.

Under such a competitive environment, port performance measurement is not only a powerful management tool for port operators, but also constitutes a most important input for informing regional and national port planning and operations. Traditionally, the performance of ports has been variously evaluated by calculating cargo-handling productivity at berth (Bendall and Stent, 1987; Tabernacle, 1995; Ashar, 1997) by measuring a single factor productivity (De Monie,1987) or by comparing actual with optimum throughput over a specific time period (Talley,1998). In recent years, significant progress has been made in the measurement of efficiency in relation to productive activities. In particular, non-parametric frontier methods such as Data Envelopment Analysis (DEA) has been developed with applications across a wide range of sectors including transit services. A recent work by De Borger, Kerstens and Costa (2002) claims that frontier models have found their way into the transport sector, and studies on the productivity and efficiency of almost all transport modes are appearing. Marlow and Paixão (2002) advocate that DEA should be used for performance measurement of ports and its suitability has been examined by Wang, Cullinane and Song (2003).

Against this background, this paper aims to provide new information on efficiency estimation by applying the techniques of DEA to the terminal data set derived from the countries' leading container ports. The paper is structured as follows: section 2 investigates performance measurement in relation to port production. A brief overview of nonparametric efficiency measurement techniques, discussing the DEA models, is included in section 3. Operationalisation and the analysis of results are provided in sections 4 and 5 respectively. Finally, conclusions are drawn in section 6.

2. PORT PRODUCTION MEASUREMENT

Performance measurement plays an important role in the development of a company or any other form of organizational Decision Making Unit (DMU). Dyson (2000) claims that performance measurement plays an essential role in evaluating production because it can define not only the current state of the system but also its future, as shown in Figure1. Performance measurement helps move the system in the desired direction through the effect exerted by the behavioral responses towards these performance measures that exist within the system. Mis-specified performance measures, however, will cause unintended consequences with the system moving in the wrong direction.



Source: Dyson (2000, p. 5)
Figure 1: Performance Measures and Organizational Development

Ports are essentially providers of service activities, in particular for vessels, cargo and inland transport. As such, it is possible that a port may provide sound service to vessel operators on the one hand and unsatisfactory service to cargo or inland transport operators on the other. Therefore, port performance cannot normally be assessed on the basis of a single value or measure. The multiple indicators of port performance can be found in the example of the Australian port industry (Talley, 1994). The indicators are selected from the perspective of the stevedore, the shipping line and the port authority (or port management). Evaluations are made by comparing indicator values for a given port over time as well as across ports for a given time period. The port performance indicators suggested by UNCTAD (1976), as shown in Table 1, underlie productivity and effectiveness measures and can be used as a reference point.

Table 1: Summary of Performance Indicators Suggested by UNCTAD

Financial indicators	Operational indicators
Tonnage worked	Arrival late
Berth occupancy revenue per ton of cargo	Waiting time
Cargo handling revenue per ton of cargo	Service time
Labour expenditure	Turn-around time
Capital equipment expenditure per ton of cargo	Tonnage per ship
Contribution per ton of cargo	Fraction of time berthed ships worked
Total contribution	Number of gangs employed per ship per shift
	Tons per ship-hour in port
	Tons per ship hour at berth
	Tons per gang hours
	Fraction of time gangs idle

Source: UNCTAD (1976, pp.7-8)

Talley (1994) goes further by attempting to build a single performance indicator – the shadow price of variable port throughput per profit dollar - to evaluate the performance of a port. This overcomes the drawback of multiple indicators, i.e. that examining whether port performance has improved or deteriorated becomes difficult when changes in some indicators improve performance and changes in others affect it negatively. In an effort to more properly evaluate port performance, several methods have been suggested, such as the estimation of a port cost function (De Neufville and Tsunokawa, 1981) the estimation of the total factor productivity of a port (Kim and Sachish, 1986) and the establishment of a port performance and efficiency model using multiple regression analysis (Tongzon, 1995).

In recent years, DEA has occasionally been used to analyze port production. Compared with traditional approaches, DEA has the advantage that consideration can be given to multiple inputs and outputs. This accords with the characteristics of port production, so that there exists, therefore, the capability of providing an overall evaluation of port performance.

3. METHODOLOGIES

3.1 Model Selection

An efficient production frontier defines the relationship between inputs and outputs by depicting graphically the maximum output obtainable from the given inputs consumed. In so doing, it reflects the current status of technology available to an industry. Ignoring all the economic complexities associated with the particular or possible source, or cause, of inefficiency (such as technical (productive), allocative or scale efficiency), at its most fundamental level, a DMU is considered efficient if it operates on the efficient frontier. On the other hand, a DMU is regarded as basically inefficient (for whatever reason) if it operates beneath the efficient production frontier.

Data Envelopment Analysis (DEA) is one of the many available alternative techniques (categorized either as econometric or as mathematical programming) for estimating an approximation to the efficient frontier. This mathematical programming techniques allow the measurement of the relative distance that an individual DMU (data observation) lies away from this estimated frontier and, thereby, also yield measures (usually in index form) of the relative inefficiency of the individual DMU in question, as compared to what amounts to an industry 'best practice' output/input ratio.

In fact, DEA is the most important non-parametric techniques to measure the efficiency of DMUs with multiple outputs and inputs. First introduced in Charnes, Cooper and Rhodes (1978), DEA has been widely used because it can be applied in a diverse variety of situations and has also been the subject of a number of theoretical extensions that have increased its flexibility, ease of use and applicability (Allen *et al*, 1997). As the counterpart of DEA, first appeared in Deprins, Simar and Tulkens (1984) and according to Lovell and Vanden Eeckaut (1993) is gradually becoming more popular.

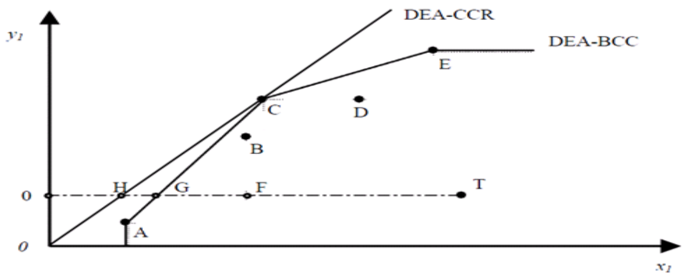


Figure 2: Non-parametric Deterministic Frontiers

DEA has its respective strengths and weaknesses (Lovell and Vanden Eeckaut, 1993). As such, a comparative study of this approach may provide greater insight into the intricacies of measuring production efficiency. Efforts in this respect include, *inter alia*, the efficiency of municipalities (Vanden Eeckaut *et al*, 1993) and the efficiency of retail banking, courts and urban transit (Tulkens, 1993).

DEA, as the deterministic non-parametric methods, assume no particular functional form for the boundary and ignore measurement error. Instead, the best practice technology is the boundary of a reconstructed production possibility set based upon directly enveloping the observations. These extremal methods use mathematical programming techniques to envelop the data (in a piecewise linear way) as tightly as possible, subject to certain production assumptions that are maintained within the mathematical programming context.

Convex non-parametric frontiers in the context of DEA allow for linear combinations of observed production units. According to this definition, all linear combinations of observations A and C are feasible in Figure 2.

Figure 2 illustrates the two most widely used DEA-models: The DEA-CCR (due to Charnes, Cooper and Rhodes, 1978) assumes constant returns to scales so that all observed production combinations can be scaled up or down proportionally. This constant return to scale DEA frontier is derived simply by the ray through the origin passing through point C. The DEA-BCC model (due to Banker, Charnes and Cooper, 1984) on the other hand, allows for variable returns to scale and is graphically represented by the piecewise linear convex frontier.

DEA-CCR and DEA-BCC models define different production possibility sets and efficiency results. As an example, the input-oriented efficiency of unit T in Figure 2 is given by the DEA-CCR model and 0G/0T by the DEA-BCC model.

3.2 Model Orientation

DEA models can be distinguished according to whether they are input- or output-oriented. One called input-oriented that aims at reducing the input amounts by as much as possible while keeping at least the present output levels, and the other, called output-oriented, maximizes output levels under at most the present input consumption.

Both orientations have their usefulness within the context of an application to the container port industry. The former is closely related to operational and managerial issues, whilst the latter is more related to port planning and strategies. A port is normally able to approximately predict its container throughput for the ensuing year at least. This is because a container port has a fairly stable customer base of shipping lines. Over the fairly short-term, container terminals should even be able to predict impending dramatic changes, such as Maersk-Sealand's decision in 2000 to move its regional hub from Singapore to the Port of Tanjung Pelepas in Malaysia. A container terminal can also attempt to predict its future throughput by studying historic data or regional economic developments. In that case, how to efficiently use the inputs is the key to saving costs in port production.

On the other hand, with rapid expansion of globalization and international trade, many container ports must frequently review their capacity in order to ensure that they can provide satisfactory services to port users and maintain their competitive edge. Sometimes, the need to build a new terminal or increase capacity is inevitable. However, before a port implements such a plan, it is of great importance for the port to know whether it has fully used its existing facilities and that input has been minimized given the output. From this perspective, the input-oriented model provides a more appropriate benchmark for the container industry. For the purposes of this study, it has been decided that input-oriented models should be chosen as the basis for the analysis. The fundamental reason for this choice is that since the main interest of the paper lies with informing policy-decisions.

3.3 Model Specification

There are two kinds of models used usually, one is DEA-CCR model, which is used to evaluate the overall efficiency (scale and technical efficiency) of the DMU, the other one is DEA-BBC, which is to evaluate the technical efficiency of the DMU only, the models for an input-oriented efficiency measurement problem are as follows.

Suppose there are n DMUs, with m input factors and S output factors; let j denote one of DMUs. The efficiency of the j^{th} DMU, with outputs Y_{ri} and (with r = 1,2,…,s) and inputs X_{ir} (with i=1,2,…,m), is calculated by the following DEA-CCR model

$$(a)\begin{cases} \min\theta \\ s.t\sum_{j=1}^{n}\theta X_{j0}-\lambda_j X_j \geq 0 \\ \sum_{j=1}^{n}\lambda_j Y_j \geq Y_{j0} \\ \lambda_j \geq 0,\ j=1,2,...,n \end{cases}$$

λ_j is the coefficients associated with the selection of an efficient frontier point for the evaluation of DMU_{j0}

θ is the efficiency of DMU_{j0}respectively.

For the model (a), there are such rules

(l)If θ=1, then the j^{th} DMU is overall efficient, that means in the system that is formed by DMUs, the output Y_{j0} has been optimum under the input X_{j0}

(2)Ifθ<1, then the j^{th} DMU is inefficient, it is scale inefficient or technical inefficient, or scale and technical inefficient. That means in the system, the output can be maintained by decreasing the input X_{j0}

(4) Suppose

$$k=\sum_{j=1}^{n}\lambda_j$$

(a) If k=1, then the DMU will be operating at constant returns to scale;
(b)If k <1, then the DMU will be operating at increasing returns to scale;
(c)If k >1, then the DMU will be operating at decreasing returns to scale.

DEA-BBC model is to add the constraint

$$\sum_{j=1}^{n} e_j\lambda_j = 1$$

that is

$$
(b)\begin{cases}
\min \theta \\
s.t\sum_{j=1}^{n} \theta X_{j0} - \lambda_j X_j \geq 0 \\
\sum_{j=1}^{n} \lambda_j Y_j \geq Y_{j0} \\
\sum_{j=1}^{n} e_j \lambda_j = 1 \\
\lambda_j \geq 0,\ j = 1, 2, \ldots, n
\end{cases}
$$

e_j is a suitably dimensioned vector of unity values.

Rules (1) and (2) for model (a) is also suitable for model (b). It is shown that the overall efficiency, calculated from the DEA-CCR m odel, can be decomposed into the technical efficien cy measured by DEA-BBC m odel and the scale efficiency. Indeed, the scale efficiency score of a DMU is the ratio of the overall efficiency to the technical efficiency, and using the DEA-BBC model can specify the major sources causing overall inefficiency.

4. OPERATIONALISTION

4.1 Definition of Variables

A thorough discussion of variable definition is provided in Cullinane, Song and Wang (2003), and can be summarized as follows. The input and output variables should reflect actual container port production as accurately as possible. T o this end, a systematic investigation of container production is necessary. As far as container port production inputs are concerned, a c ontainer terminal depends crucially on the efficient use of labour, land and equipment. T he total quay length, the terminal area, t he number of gantry cranes, the number of yard gantry cranes and the number of straddle carriers are the most suitable to be incorporated into the models as the input variables. In the light of the unavailability or unreliability of direct data, information on labour inputs is derived from a p re determined relationship to term inal facilities. On the other hand, container throughput is unquestionably the most important and wi dely accepted indicator of port or terminal output. Almost all the previous studies treat it as an output variable, because it closely relates to the need for cargo -related facilities an d serv ices and is th e primary b asis u pon wh ich container ports are co mpared, especially in assessing th eir relativ e size, in vestment mag nitude or activity lev els. An other con sideration is that container throughput is the most appropriate and analytically tractable indicator of the effectiveness of the production of a port.

4.2 Data Sources

Data were collected from various sources to test and run t he model as form ulated in the previous section. The data are concerned with two output measures (Throughputs, Vessel Calls), three input measures (Number of berths, Number of cranes, Quay lengths) of six countries (Philippines, Brunei Darussalam, Indonesia, Myanmar, Singapore, China). The six countries' sea ports were sel ected and a ggregated o n t he basi s of i ndustrial co mplex, po pulation m ap, t ransportation network, province and metropolitan and special city size with the help of an expert's advice in transportation as follows with each region numbered with its na me: Philippines seaports, Brunei seaports, Indonesia seaports, Myan mar seaports, Singapore seaports and China seaports.

Table-2: Input and Output Factor of Ports

Country	Philippines	Brunei	Indonesia	Myanmar	Singapore	China
Number of cranes	18	10	60	20	140	3970
Number of berths	14	9	18	10	54	3659
Quay length(meter)	4914.5	350 600 400			1600	807102
Through put(million tons)	392.86	10.78 564.13		110.98	2078.24	33160
Vessel calls	89057	11252	506466	10119	1235068	11869200

Source: Ministry of Transport of P.R.C, Research in China, Brunei Darussalam website
Source: People Republic of China, China Port Industry report, 2006-2007 & 2008-2009

5. RESULTS OF THE EFFICIENCY ANALYSIS

In this section, the DEA linear programming model is validated using the data given in Section 3 to check whether or not the model represents the real situations in China and ASEAN-5. In this paper, we use MATLAB software to do the calculation (programs are at the rare of the paper), the results are in the table 3.

Table-3: Port Efficiency of CCR and BCC Model

Country No.	CCR BBC			$\sum_{j=1}^{n}\lambda_j$	RTS	Scale score
	θ (Overall Efficiencies)	Reference	θ (Technical Efficiencies)			
c_1	1	c_1	1 1		Constant	1
c_2	0.1257	c_5	0.4482 0.0091		Increase	0.2805
c_3	1	c_3	1 1		Constant	1
c_4	0.3681	c_1, c_5	0.3738 0.0643		Increase	0.9848
c_5	1	c_5	1 1		Constant	1
c_6	0.8038	c_1, c_3, c_5, c_6	1 34.758		Decrease	0.8038
Average 0.7163		0.8037				0.8449

$c_1 = Philippines, c_2 = Brunei, c_3 = Indonesia, c_4 = Myanmar, c_5 = Singapore, c_6 = China$

5.1 Overall Efficiency Analysis

The CCR results, listed in Column 2 of Table 3, show that the ports of Philippines, Indonesia and Singapore performed the best ports than the other ports of countries when evaluated on the constant returns-to-scale assumption associated with this model, as evidenced by the fact that ports of Brunei and Myanmar were below the average and ports of China were above the average. Ports of Philippines, Indonesia and Singapore are the best performers. However, ports of Singapore are the best performers among the ports of 3 countries, and furthermore it is the ports of country most frequently referenced for evaluating inefficient ports of countries. It is used as a reference for all inefficient ports of countries and serves as the most influential referent, i.e., with the largest λ value. Ports of Singapore are the most efficient one in CCR measure. For confirmation, we might note that ports of Singapore are famous for its unique managerial strategies under the strong leadership of its owner.

5.2 Scale Efficiency and Technical Efficiency Analysis

The BCC scores provide efficiency evaluations using a local measure of scale, i.e. under variable returns-to-scale. In this model, ports of China are accorded efficient status in addition to the three CCR efficient ports of countries — which retain their previous efficient status. Ports of China's full efficiency with the BCC model is caused by its use of the smallest amount of inputs. Ports of Brunei and Myanmar are below average.

5.3 Scale Efficiency

The scale efficiency as defined by the ratio, CCR/BCC, exhibits large differences between the ports of countries. Ports of China and Brunei are below average whereas other ports of countries are above it. This may mean Ports of Philippines, Indonesia, Myanmar and Singapore are in an advantageous condition compared with those in the China and Brunei. But the ports of Brunei are the worst among the ports of 6 countries. Their overall inefficiencies (CCR) are mainly attributed to their inefficient operations or management.

Ports of countries with full efficiency in the CCR score are also efficient in the BCC model, the region where constant returns-to-scale prevails. Ports of Philippines, Indonesia and Singapore have this status. Ports of China and Myanmar show that they have a possibility to improve their efficiency by scaling up their activities. This observation leads us to hypothetically study the desirability of merging low ranked ports of China and Myanmar.

6. CONCLUSION

In this paper, DEA models are applied to evaluate the efficiency of performance of ports in different countries. Two DEA models (CCR model and BBC model) are used to evaluate the overall efficiency, technical efficiency, and scale efficiency of all ports of countries. Based on the results, the specific directions for the inefficient ports of countries to improve their operation efficiencies are discussed. In addition, the input and output factors for the performance of ports of countries are proposed, and successive period of one country can be taken as the DMUs to do the evaluation by DEA, which will help the owner to identify its improvement. Therefore, this study may provide useful information for the policy makers to implement ports performance better.

ACKNOWLEDGEMENT

This work was supported in part by Shanghai Science Commission Project (No. 09DZ2250400), Shanghai Science Commission International Cooperation Project (No. 09 530708200), and Shanghai Municipal Science Commission Local University Capability Project (No. 08 170511300), Shanghai Education Commission Leading Academic Discipline Project (No. J50604) and Shanghai Science Commission Project (No. 10692103500).

7. REFERENCES

Ali, A. I. and Seiford, L. M. (1993), The Mathematical Programming Approach to Efficiency Analysis, In Fried, H. Lovell C.A. K. and

Schmidt, S. (eds.), The Measurement of Productive Efficiency: Techniques and Applications, Oxford: Oxford University Press, pp. 160-194.

Cooper, W.W., Seiford, L. M. and Tone, K. (2000), Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software, Kluwer Academic Publishers: Boston.

Vanden Eeckaut, P., Tulkens H. and Jamar M.A.(1993), Cost Efficiency in Belgian Municipalities, in Fried, H.O., Lovell, C.A.K., and Schmidt, S.S.(eds.), The Measurement of Productive Efficiency, New York: Oxford University Press.

Wang, T., Cullinane, K.P.B. and Song, D-W. (2003), The Efficiency of Container Ports and Terminals: Assessing the Role of Data Envelopment Analysis as a Measurement Methodology, in Lee, T-W. and Cullinane, K. (eds.), Maritime Transport in Asia, Ashgate Publishing, Aldershot, UK. Forthcoming.

Charnes, A., Cooper, W. W. and Rhodes, E. (1978), Measuring the Efficiency of Decision Making Units, European Journal of Operational Research, Vol. 2, pp. 429-444.

Cullinane, K.P.B., Song, D-W. & Wang, T. (2003), A Comparison of Mathematical Programming Approaches to Estimating Container Port Production Efficiency, Journal of Productivity Analysis, (under review).

De Borger, B., Kerstens, K. Costa, A. (2002), Public Transit Performance: What does One Learn from Frontier Studies, Transport Reviews, Vol. 22, No. 1, pp. 1-38.

De Neufville, R. and Tsunokawa, K. (1981) Productivity and Returns to Scale of Container Ports, Maritime Policy and Management, Vol. 8, No. 2, pp. 121-129.

Dyson, R. G (2000), Performance Measurement and Data Envelopment Analysis – Ranking are ranks! OR Insight, Vol. 13, No. 4, pp 3-8.

Epstein, M. K. and Henderson, J. C. (1989), Data Envelopment Analysis for Managerial Control and Diagnosis, Decision Sciences, Vol. 20, No. 1, pp. 90-119.

Kim, M. and Sachish, A. (1986), The Structure of Production, Technical Change and Productivity in a Port, Journal of Industrial Economics, Vol. 35, No. 2, pp. 209-223.

Martinez-Budria, E., Diaz -Armas, R., Navarro-Ibanez, M. and Ravelo-Mesa, T. (1999) A Study of the Efficiency of Spanish Port Authorities Using Data Envelopment Analysis, International Journal of Transport Economics, Vol. XXVI, No. 2, pp. 237-253.

Kim, M. and Sachish, A. (1986), The Structure of Production, Technical Change and Productivity in a Port, Journal of Roll, Y. and Hayuth, Y. (1993) Port Performance Comparison Applying Data Envelopment Analysis (DEA), Maritime Policy and Management, Vol. 20, No. 2, pp. 153-161.

Seiford, L. M. and Thrall, R. (1990), Recent Development in DEA: the Mathematical Programming Approach to Frontier Analysis, Journal of econometrics, Vol. 46, Vol. ½ (October/November), pp. 7-38.

Tabernacle, J. B. (1995), A Study of the Changes in Performance of Quayside Container Cranes, Maritime Policy and Teng- Management, Vol. 22, No. 2, pp. 115-124.

Fei Wang, Dr Dong-Wook Song, Prof. Kevin Cullinane.(2003) ,Container Port Production Efficiency: A Vol.5, Comparative Study of DEA and FDH Approaches Journal of the Eastern Asia Society for Transportation Studies, October, ,pp.698-713.

Session D1: SCM 4

·**Day1: Sep. 15 (Wed.)**

·**Time: 16:40 - 18:00**

·**Chair: Kwangyeol Ryu**

·**Room: Camellia, 5F**

LOGISTICAL SERVICE PROVIDER TRIADIC RELATIONSHIPS

Paul Childerhouse[1*], Hyung Jun Ahn[1], Habin Lee[3], Wen Luo[1], and Gottfried Vossen[4]

[1*]Corresponding author, Department of Management Systems
Waikato Management School
University of Waikato
Hamilton 3240
New Zealand
e-mail: pchilder@waikato.ac.nz

[2]College of Business Administration
Hongik University
Seoul, Korea

[3]Brunel Business School, Brunel University
Uxbridge, Middlesex
UK

[4]School of Business Administration and Economics
University of Muenster
48149 Muenster
Germany

Abstract: The concept of supply chain management has evolved from focussing initially on functional co-ordination within an organisation, then to external dyadic integration with suppliers and customers and more recently towards a holistic network perspective. A triad is the simplest meaningful sub-element of a network, and as such will be used as the unit of analysis for this research into relationship continuity and information connectivity of logistical service providers and their associated supply chains.

The type of interdependence between network players dictates the type of coordination required. Most organisations operate in supply chains with sequential or reciprocal interdependence, both of which require advanced forms of information systems enabled coordination. Networks are dynamic, the connections, working relationships and interdependency evolve over time. Thus any information system solution needs to be reasonably robust to alternative network dynamics. In this paper, we match information systems to supply chain triadic scenarios in an attempt to ensure appropriate connectively.

1. INTRODUCTION

This paper investigates inter-firm interactions from a sociocultural perspective. In particular, we examine the relationship dynamics of a network of inter-connected firms with shared end consumers. The social network perspective has gained significant momentum in the management literature (Wang, Heng and Chau, 2007). In this paper we use the psychological concept of balanced theory (Simmel, 1950; Heider, 1958) to make sense of the dynamic inter-relationships of an organisation's supply chain network. The most important dimensions of change in business networks concerns the development of activity links, resources ties, and actor relationships bonds (Gadde and Hakansson, 2001).

A triad is the smallest meaningful sub-set of a network (Madhavan, Gnyawali and He, 2004) and as such will be used as the unit of analysis throughout this paper. Figure 1 is a simplistic representation of the multi-layered complex business interactions that make up supply chain networks. The actors are represented by nodes and the connections between them as links. A triadic sub-set of the entire network is illustrated as the grey shaded area in Figure 1. Three actors, 'A', 'B' and 'C' are highlighted and their three links, 'A' with 'B', 'A' with 'C' and 'B' with 'C'. Each actor also has a potential mediating role in the relationship between the other two as indicated by the dashed arrow from actor 'A' to the link between 'B' and 'C'. Thus we contend that a representative sub-set of a network can be investigated via triads. This cannot be said for dyads that overly simplify the social complexities of real world business interactions.

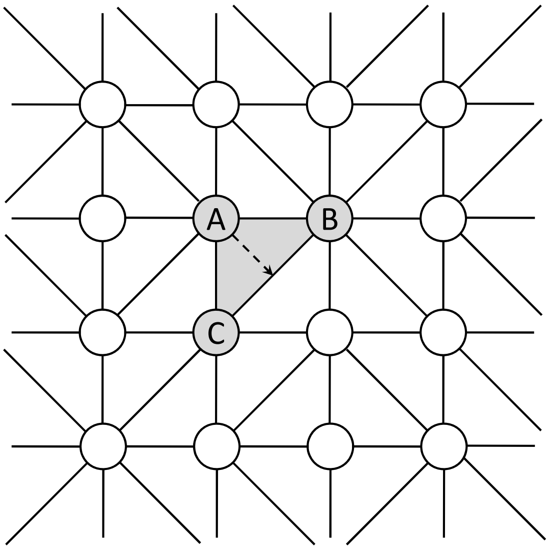


Figure 1. The Underlying Triadic Structures that Combine to form Complex Supply China Networks

Balanced theory (Simmel, 1950; Heider, 1958) has been developed in psychology to understand the inter-relationships between individuals. The resultant insight identifies four balanced triads and four that are unbalanced. It is argued that over time certain combinations of positive and negative relationships either remain stable or evolve to one of the alternative possible triadic states. The aim of this research is to use this social perspective on the interactions of individuals to investigate business triads consisting of a logistics service provider (LSP) and their customers and suppliers. Once we have established the alternative triadic states we will then go on to explore the communication (links) mediums between each actor, by discussing ways to match information system connectively with the reported eight triadic states. A brief overview of the relationship management literature follows. This is complemented by a review of balanced theory and the identification of eight alternative states.

2. INTER-ORGANIZATIONAL RELATIONSHIP MANAGEMENT

Table 1 describes and compares five leading theoretical lenses used to study inter-organizational relationships. From the table, it can be concluded that balance theory is arguably the more suitable approach to investigate triadic relationships. This theoretical approach has already been shown to be applicable to larger social groups such as of organizations (Madhavan, Gnyawali and He 2004).

Table 1. Conceptual Evolution of Supply Chain Relationship Management

Theory Descriptions		Representative literature
Transaction Costs Analysis (TCA)	Focuses on how an organization should organize its relationship to minimize its own costs. The focus is on dyadic relationships between the focal organization and one other actor.	Dyer (1997); Madhok & Tallman (1998)
Resource Dependency Theory (RDT)	Collaborative organizations should engage with each other to share unique resources. Typically dyads are used to demonstrate the advantages of partnering.	Bartholomew (1997); Das & Teng (1998);
Institutional Theory (IT)	This theory suggests that institutional environments could bring more dynamics to inter-organizational relationship. Once again a dyadic relationship view is dominant.	Baum & Oliver, 1991 ; Kraatz (1998);

Network Theory (NT)	Every organization needs to cooperate with its direct partners and the partners' partners. An actor's place within this complex network is considered. Each firm has direct and/or indirect relationships with all the other actors. An overall network view is proposed.	Oliver, 1990 , (Haakansson and Ford, 2002) (Halldorsson, 2002)
Balance Theory (BT)	Conceptualizes the motives and drivers toward psychological balance. The focus is on the triadic balance between three parties. How the three nodes and resultant three links influence each other and the third party.	Heider 1958;Madhavan, Gnyawali and He 2004

3. BALANCED THEORY

Balance theory has been developed from behavioural psychology. It has been used to investigate relationships among individuals for more than thirty years (Davis and Leinhardt 1972; Alessio 1990). Balance theory is the only representative theory that addresses triadic relationship explicitly. According to balance theory, there are two general relationships between individual and or organizations: positive and negative. Typically, a positive relationship presents a cooperative, mutual trust, and collaborative relationship between two firms (Morgan and Hunt 1994; Uzzi 1997). A negative relationship presents an adversarial, distrust and transactional cost focus between two firms (Johnston et al. 2004; Griffith, Harvey and Lusch 2006). Typically researches adopt "+" and "-" as signs for positive and negative relationships (Andersson-Cederholm & Gyimothyb 2010; Choi & Wu, 2009). Furthermore, the different combination of positive and negative relationships between three relevant people or parties make four balanced triadic states and four unbalance triadic states.

Transactional relationships are akin to adversarial, exit-based relationships with low trust and minimal mutual commitment. The output these kinds of relationships tend to be win-lose: one side takes more of the profit at the expense of the other. In a balanced theory sense these win-lose adversarial (Johnston et al. 2004; Griffith, Harvey and Lusch 2006) relationships can be viewed as "negative" ties between two firms (Andersson-Cederholm & Gyimothyb 2010). On the contrary, collaboration represents a highly cooperative partnership between two organizations with high levels of mutual trust and involvement. This collaborative approach can help both actors develop win-win scenarios. Thus we can call this type of link between actors "positive" (Morgan and Hunt 1994; Uzzi 1997; Phillips, Liu, & Costello, 1998; Choi & Wu, 2009) in a balanced theory context. To date no supply chain research has been conducted using balanced theory to make sense of business networks and more specifically triads. Hence our original contribution is primarily conceptual via cross-disciplinary application of a theoretical lens.

4. EIGHT TRIADIC SCENARIOS

Figure 2 provides a visual representation of the eight feasible triadic scenarios with binomial relationship links. On the left hand side is the transactional triad, here the three actors (nodes) are loosely connected (represented by their close proximity to one another) via short term transactions. The next form of triads are the partnerships, here one strong collaborative bound (represented by a link) exists between two of the three actors. There are three alternative partnership triads, obviously based around the alternative combination of pairs. Collaborative triads are the third type of triads illustrated in figure 2. In this scenario one actor is collaborating with the two other actors, but these two are themselves adversarial towards each other. Once more three alterative combinations are feasible, this time based on which actor in the triad is the duel collaborator. The final triadic scenario is the cluster when all three actors collaborate with each other. Simmel (1950), Heider (1958) and more recently Hummon & Doreian (2003) argue that the transactional and collaborative scenarios are unbalanced and will not last in the long term, whereas the partnership and cluster triads are more balanced and hence should provide a business with stable inter-organisational relationships. Each of these alternative triadic states will now be explored.

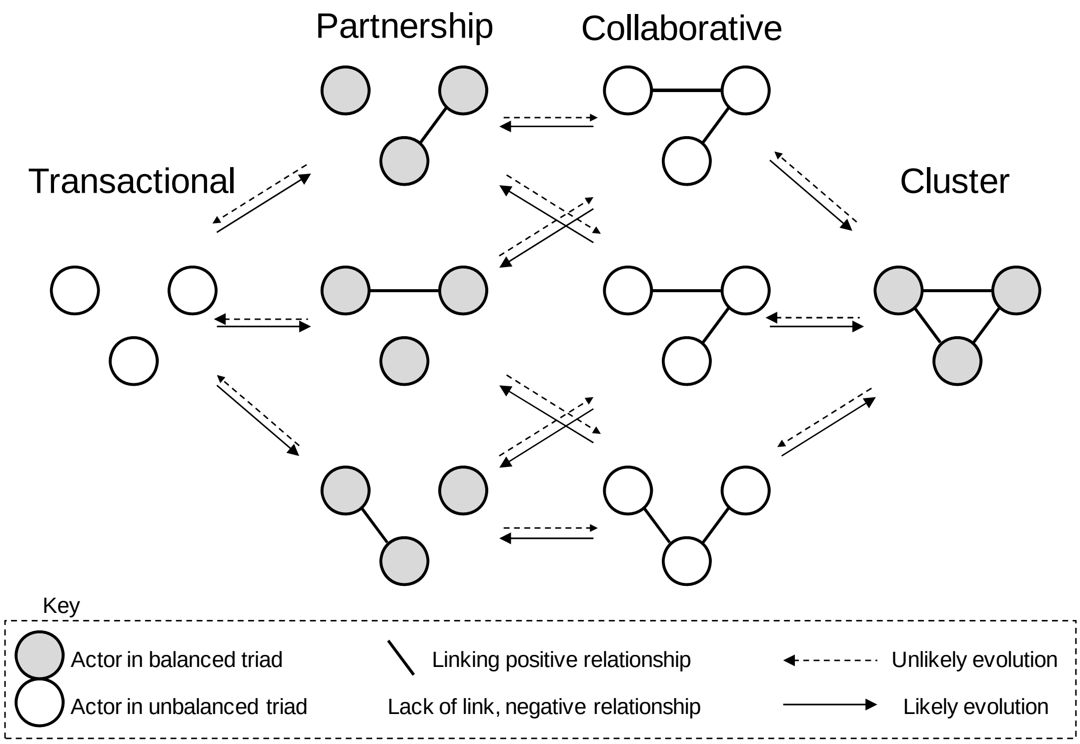


Figure 2. Eight Triadic Scenarios

4.1 Transactional

According to Heider's principles (1958), when every actor only has negative links with the other two, the triad will be unbalance and will be not last. This is because all three actors only perceive tension with others in the triad. If they want the triad to endure they should attempt to transform one of the negative links to positive to stabilize the triadic structure (Heider 1958; Alessio, 1990) and move to a partnership triad. When all of the links are adversarial in a triadic situation, it is not hard to see that the triad would dissolve if none of the actors wants the relationship to endure. This is very common in the real world, with all three actors taking an adversarial posture toward one another and attempting to maximize their own gain, often at the expense of the others.

4.2 Partnership

Enhancement in overall supply chain performance is a well documented outcomes of partnering initiatives (Jitpaiboon, Dangol & Walters, 2009; Zailani & Rajagopal, 2005; Flynn, Huo & Zhao, 2010). As mentioned in Heider and Alessio's (1990) research findings, when two indivuals have common negative sentiments to a third person, these two will develop and strengthen their positive sentiments to each other (Heidr, 1958; Alessio, 1990) i.e. your enemy is my enemy, so let's be friends. These partnerships are thus stable, but the third actor may be replaced overtime because they will be on the outside and may become concerned that the other two actors are scheming against them.

4.3 Collaborative

As discussed in Heider's balance theory and Simmel's model, when both 'B' and 'C' like 'A', the triadic structure will be unstable if 'B' and 'C' dislike each other (Heider, 1958; Simmel, 1950; Hummon & Doreian, 2003). Thus this kind of triadic structure will be dismantled reasonably quickly because either 'A' or 'B' will remove themselves from the triad (Heider, 1958; Hummon & Doreian, 2003). As illustrated by the arrows in Figure 2 these triads will either evolve back towards a partnership by the removal of a link or alternatively by adding a link to form a cluster. In essence the actor with both collaborative links holds the power in this triad, the other two realise this and will therefore attempt to readdress this power imbalance.

4.4 Cluster

Just as in dyadic relationships, cluster organizations treat partners equitably based on mutual trust and co-dependency. This facilitates integrative and collaborative processes and provides continuity for collaboration (Leger, Cassivi, Hadaya & Cya, 2006). As discussed in Simmel's (1950) triadic research, if individual 'A' and 'B' know 'C' treats them in a different way, 'A' and 'B' will resist getting any closer to each other. Only after 'C' treats both 'A' and 'B' equally, will 'A' and 'B' be glad to get together (Simmel, 1950). Furthermore, according to Choi & Wu's (2009) research, in this kind of triadic structure, the customer is responsible for solving conflicts and building a positive culture that prohibits opportunistic behaviour. This kind of triadic relationship should last for the long term because all of the three partners collaboratively work together and share the benefits in a win-win-win situation.

5. PRACTICAL APPLICATION OF TRIAD RELATIONSHIP DYNAMICS

Many organisations work closely with two other actors to achieve a shared goal. Figure 3 illustrates four commonplace real world scenarios that contain clearly definable triads. On the far left is an ideal scenario for a producer with duel customers (note the arrows represent the major direction in which material is flowing). Arguable the supplier has the power if they are producing a unique or innovative product (ref). The supplier can them chose to develop collaborative links with neither, either or both customers (translational, partnership or collaborative triads). Balanced theory would suggest that if the supplier builds collaborative links with both customers (thus forming a collaborative triad) this will not last in the long term as these two customers will fear that one is getting preferential treatment over the other.

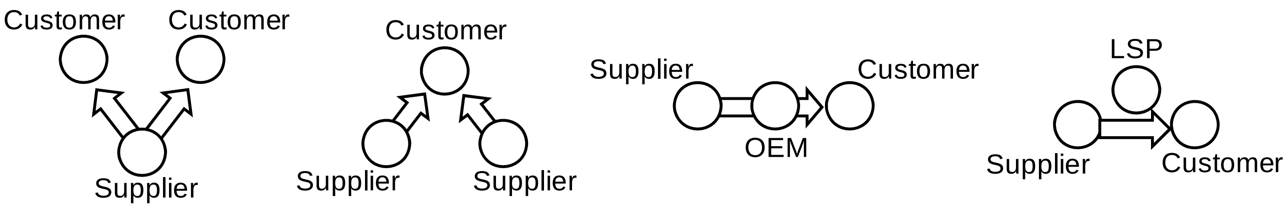


Figure 3. Four Practical Supply chain Triad Examples

The second practical triad scenario illustrated in Figure 3 is duel sourcing. In this setting two suppliers are providing similar products and/or services to a single customer. Here the customer has the opportunity to trade the suppliers off one another via transactional relationships or build a partnership with one or both suppliers. The latter option is once more problematic as it will result in each supplier becoming concerned that customer is giving their competitor preferential treatment.

The third supply chain setting illustrated in Figure 3 relates to specified suppliers. Here the customer not only has a relationship with their direct supplier (OEM) but also the supplier's supplier. Arguably it is in the interests of the OEM to build a collaborative triad by linking closely with both the customer and specified supplier. However, the customer determines who this second tier supplier should be and thus dictates the ways of working. Two common approaches result, a cluster triad with all actors collaborating with each other or a partnership between the customer and either the OEM or supplier. In the first instance the customer may play a mediation role to bring the other two parties together. In the later instance the customer may team up with one actor in order to jointly extract as much value as possible from the other. The fourth real world setting for this triad research evolves third party logistics providers.

Illustrated on the right hand side of Figure 3 is the position Logistics Service Providers (LSP) often find themselves. The flow of goods is represented by the arrow from the supplier (producer) to the customer (user). The LSP is brought into these triads to link the suppliers and customers together. The relationship dynamics of each of the alternative eight triadic states will now be explored within this context.

5.1 Transactional

Arguably the fully transactional triad is the most commonplace in practice. Here the LSP is contracted by a supplier to perform a logistical service in order to link this supplier with a customer. Despite many years of research promoting the advantages of collaboration most organisations still act in a transactional manner with nearly all of their customers, suppliers and sub-contractors (ref). These transactional triads are constantly evolving and changing (actors leave and new ones join), they are inherently unbalanced. The obvious advancement is a formation of a single link between any of these partners, thus giving rise to three alternative partnership balanced triads.

5.2 Partnership

The type of role played by the LSP now becomes central, are they part of the partnership or on the outside. If they develop a close relationship with either the customer or more likely the supplier then they would have developed a more balanced scenario with themselves as a pivotal actor. However if the supplier and customer form a partnership first they will be marginalized. This type of triad is balanced for the dyadic partnership, however there is somewhat of a revolving door for the actor on the outside. As such this outsider may attempt to build a close relationship with either of the other two and hence develop a collaborative triad.

5.3 Collaborative

There are three co llaborative triads, dependent on which one of the three actors is the central collaborator. A LSP can build close relationships with both the customer and supplier. Alternatively the LSP could have one collaborative link and one transactional with either the customer or supplier. In all these cases the relationships are unbalanced and arguable will not last in to the medium term. In the first case the customer and supplier will become concerned that the LSP is trading them off one another and in essence has too much influence. If the LSP is on the outer they themselves will be concerned that there collaborator is treating the other actor better than themselves, once more leading to the erosion of trust and the dissolvent of the triad. Two alternative directions for change are the breakup of one of the links due to jealousy thus becoming a partnership or the through the development of collaborative behaviour between the two transactional actors thus forming a cluster.

5.4 Cluster

The final triadic scenario is one which contains only collaborative links. Thus the three actors have formed a stable cluster. Here the LSP operates in an open and collaborative manner with both the customer and supplier and these two also act positively towards one another. This should result in a win-win-win scenario for all actors as they combine their core strengths to collectively compete. If for some reason any one of the actors circumstances change so that collaboration is no longer pursued this will initial result in a collaborative triad and as this is unstable eventually a partnership triad with the other two actors retaining their hold on power.

6. INFORMATION SYSTEMS SYMMETRY

Based on our modelling of the triadic relationships, Figure 4 shows the mapping between different types of information systems and each triad state, along with a continuum about the degree to which each state is collaborative or transactional. Supply chain portals can best support the relationship characteristics of clusters, whereas B2B exchanges or auctions can be used for transactional triads.

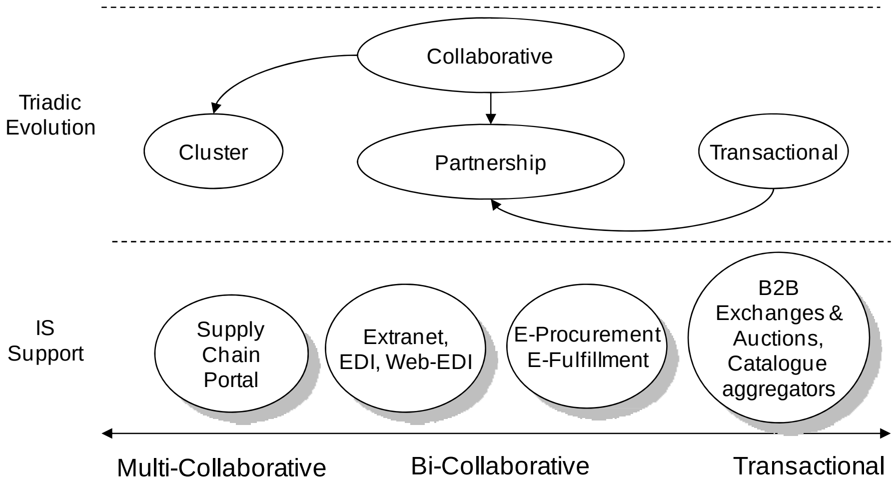


Figure 4. Triads and Supply Chain Information Systems

We can expect the adoption of the information systems suitable for the balanced states facilitate the proposed evolution from unbalance states to balance ones. For example, adoption of Extranet, EDI, and e-procurement systems can facilitate the transition from transactional to partnership triads. Building a supply chain portal can similarly help companies transit from collaborative to cluster triads. Therefore, our model and propositions can also be useful for helping organisational select and adopt suitable information systems to stabilize their supply chain relationships in the long term.

7. CONCLUSIONS

The conceptual exploration of relationship dynamics between three organisations has provided valuable insight into the potential stability and evolution of supply chain networks. The role of a mediating actor bringing other actors together requires considerable more research. Our next steps are to fully explore this new perspective on social networks through case based comparisons of a multi-national sample. The references are available on request.

PRODUCTION AND PROCUREMENT PLANS FOR THE DIVERSIFICATION OF SUPPLY NETWORK RISKS

Seungjin Oh[1], Kwangyeol Ryu[2*], Hyunbo Cho[1], and Mooyoung Jung[3*]

[1]Department of Industrial and Management Engineering
Pohang University of Science and Technology (POSTECH)
Hyoja san 31, Pohang 790-784, Korea
e-mail: halmea@postech.ac.kr

[2]Department of Industrial Engineering
Pusan National University
Geumjeong-gu, Busan 609-735, Korea
e-mail: kyryu@pusan.ac.kr

[3]School of Technology Management
Ulsan National Institute of Science and Technology (UNIST)
Banyeon-ri 100, Ulsan 689-798, Korea
e-mail: myjung@unist.ac.kr

Abstract: A major concern of many global enterprises is to diversify their supply network risks. While many strategies for mitigating supply network risks have been proposed, it still remains unclear how to apply the strategies to a supply network in operational perspective. This paper examines how to plan the production and procurement of supply network for the diversification of risks. In this paper, a framework for generating the production and procurement plan based on the type of risks is proposed. A methodology for evaluating the degree of diversification of supply network according to the risks is also presented. The degree of diversification is evaluated using fuzzy techniques. The input parameters of fuzzy functions are specified with mathematical models. To illustrate the proposed framework and methodology, numerical examples are presented.

1. INTRODUCTION

Supply network management (supply chain management) is one of the most important issues in improving enterprises' financial performance. Many different ways of optimizing supply networks have been proposed so that enterprises can improve their financial performance. In the 1990s, most of the research endeavors have focused on static and centralized supply networks because past business processes were simple and static (Burnes and New, 1996). In recent global market places, however, there are many supply network risks such as supply disruptions (Oke and Gopalakrishnan, 2009), international regulations, ever changing demand (Kleindorfer and Saad, 2005) and a technology transfer issue. Dell case is a famous example of significant supply disruptions. In 2006, due to a fire hazard, Dell recalled 4 million laptop computer batteries made by Sony. The supply network risks in companies, which adopt the static supply network, make it difficult to optimize their performance. To diversify the risks and dynamics of market places, recent trends regarding globalization emphasize flexible supply network management. Flexible management concept has proposed many strategic practices such as total quality management (TQM) (Carter et al., 1998), postponement (Su et al., 2005), Just-In-Time (Wildemann and Carlson, 1987), flexible manufacturing process, multiple suppliers and flexible supply contracts to mitigate supply chain risks (Tang and Tomlin, 2008). If Dell applied the multiple supply strategies to their global supply networks, they could diversify risks about supply disruptions. Many recent researches proposed various strategies for the diversification of risk, however, it remains unclear how to apply those strategies to the supply network in operational perspective. In a case where decision makers select the multiple supply strategy, the optimized number of suppliers and procurement plans of the suppliers should be determined in production (or procurement) planning phase.

The main objective of this paper is to develop a framework for evaluating the degree of diversification of supply network according to the risks, and for generating the production and procurement plan based on the evaluated degree of diversification. To develop the framework, strategies related to the mitigating of supply network risks are identified, and a method is proposed for evaluating the degree of diversification of a supply network. The method adopts a fuzzy inference system to evaluate the diversification of the supply network in terms of numerical value. Input parameters of the fuzzy inference system are specified with mathematical models. To generate the production and procurement plan, the objective function and decision variables are also formulated.

The rest of this paper is organized as follows. In section 2, the related works including strategies for mitigating risks are surveyed. The framework including the methodology for evaluating the degree of diversification and the objective function for generating the production and procurement plan is presented in section 3. A numerical example for an exemplary global supply network is presented in section 4 to illustrate the proposed methodology. Section 5 concludes the paper.

2. RELATED WORKS

Many researchers have focused on the management and mitigation of the supply network risks (Christopher and Lee, 2004; Chopra and Sodhi, 2004). Demeter et al. (2006) have examined the relationship between focal company strategies and supply network configuration based on interviews in the Hungarian automotive industry. Tang and Tomlin (2008) examined the risk factors related with supply network management and suggested a way of mitigating supply network risks. They focused on the flexibility of supply chain for risk management. Three major types of supply chain risks that occur regularly are defined. Table 1 summarizes the six major types of risks and strategies for diversifying risks.

Table 1. Three major types of risks and related strategies

Risk	Strategies
Supply risks (Supply disruption risks, Supply cost risks)	Flexible procurement, Flexible delivery
Process risks	Flexible manufacturing, Flexible delivery
Demand risks	Postponement, Flexible manufacturing, Flexible delivery

In this paper, we focus on two types of strategies, namely, flexible procurement and flexible manufacturing. Flexible procurement strategy is used to diversify the risks such as supply disruptions, supply cost risk and uncertain demand from manufacturers. Flexible manufacturing strategy is usually applied to uncertain customer demand risks and fluctuation of process capacity risks. Flexible delivery is a principal strategy for flexible procurement and manufacturing strategy because it is impossible to implement the flexibility strategies without a flexible delivery system. These flexibility strategies reconfigured the supply network structure and production (or procurement) plans for mitigating the supply network risks. Table 2 summarizes the relationship between flexibility strategies and configuration of supply network structure.

Table 2. Configuration types according to strategies

Strategy	Configuration
Flexible procurement	Multiple suppliers
Flexible manufacturing	Multiple plants, Multi-product manufacturing

Flexible procurement strategy reconfigures the supply network to a multiple suppliers' structure. In a configuration of multiple suppliers, risks related with suppliers can be dispersed because of the increased number of suppliers. Multi-product manufacturing type shifts production quantities across internal resources through the increasing variety of products made in plants. Multiple plants shift quantities across external resources (new factories or plants).

The degree of diversification of supply networks should be evaluated according to the supply network risks. For example, let us assume a supply network which has ten plants and one supplier. The degree of diversification of this supply network may be very good for diversifying the ever changing demand risks because it has many plants and can distribute production on diverse plants; in contrast, the degree of diversification may be very poor from the perspective of supply disruption risks. In the next section, a framework for evaluating the degree of diversification of supply network according to the supply network risks including the supply disruption risk, the process risk, and the demand risk will be presented.

3. FRAMEWORK FOR GENERATING PRODUCTION AND PRODUCREMENT PLANS

A framework is composed of a supply network architecture, a degree of diversification evaluation model, an objective function for generating the production and procurement plan. An architecture of supply network is composed of nodes, relationships and environment (Harland et al. 2001). As illustrated in Figure 1, the basic architecture of supply network shows a hierarchical structure, and a node (N) represents an assembler, a manufacturer or a supplier. The contract between two nodes is represented by a linkage (R). A business opportunity and customer requirements arise in environment (E) and are detected by nodes.

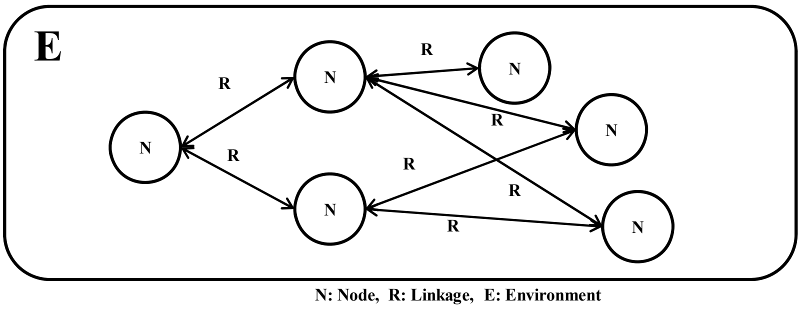


Figure 1. Hierarchical supply network

To diversify the supply network risks in operational perspective, the degree of diversification of supply network, which is referred to as *DoD*, corresponding to the risk should be represented in numerical value. *DoD* means how the structure and the production (or procurement) plan of the supply network are suitable for diversifying the risk. The formula and input data in the *DoD* evaluation model differ depending on the type of risks. In the supplier disruption risk case, the optimized number of suppliers and procurement plans of the suppliers are used to evaluate the *DoD*.

3.1 Objective Function for Generating the Production and Procurement plans

There are many researches about the supply network optimization models (including objective functions and constraints) and methodologies (such as integer programming, heuristics and simulations) regarding the profit maximization. General optimization models using profit maximization are subject to constraints imposed on customers' demand, suppliers' capacity, production capacity of plants, etc. In risk management, however, optimization through profit maximization is not suitable because there are many strategic constraints which are harmful to the profit of supply networks in operational perspective.

The proposed objective function is subject to constraints imposed on *DoD* as well as typical constraints. Constraints imposed on *DoD* lead the supply networks to diversify the production and procurement on diverse plants and suppliers. In proposed objective function and constraints, the following notations are used.

g_N	goal of node *N*
i	set of plants (i = 1,…, N)
j	set of suppliers (j = 1,…, M)
k	set of customers (k = 1,…, K)
l	set of products (l = 1,…, L)
$\underline{u}$	set of parts (u = 1,…, U)
n	index of *DoD* (n = 1, 2, 3)
v_{kl}	unit price of customer k for product l
As_{il}	unit processing cost for product l by plant i
t_{ikl}	unit transportation cost for product l from plant i to customer k
Hs_{il}	unit inventory cost for product l in plant i
P_{iju}	unit purchasing cost of plant i for part u from supplier j
H_{iu}	unit inventory cost for part u in plant i
d_{kl}	demand of customer j for product l
c_{il}	capacity of plant i for product l
c_i	total capacity of plant i

D_{iu} demand of plant *i* for part *u*
C_{ju} capacity of supplier *j* for part *u*

The following decision variables are used.

x_{ikl} = quantity of product l made in plant i for customer k
y_{iju} = quantity of part u purchased from supplier j by plant i

Proposed objective function and constraints are as follows.

$$\max g_N = \quad = \sum_{k=1}^{K}\sum_{l=1}^{L} v_{kl} d_{kl} - \sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{l=1}^{L}\left(As_{il} + t_{ikl} + \frac{1}{2}Hs_{il}\right)x_{ikl} - \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{u=1}^{U}\left(P_{iju} + \frac{1}{2}H_{iu}\right)y_{iju} \quad (1)$$

$$\text{subject to} \quad \sum_{i=1}^{N} x_{ikl} \geq d_{kl} \quad \forall k \text{ and } \forall l \quad (2)$$

$$\sum_{k=1}^{K} x_{ikl} \leq c_{il} \quad \forall i \text{ and } \forall l \quad (3)$$

$$\sum_{j=1}^{M} y_{iju} \geq D_{iu} \quad \forall i \text{ and } \forall u \quad (4)$$

$$\sum_{i=1}^{N} y_{iju} \leq C_{ju} \quad \forall j \text{ and } \forall u \quad (5)$$

$$\sum_{k=1}^{K}\sum_{l=1}^{L} x_{ikl} \leq c_i \quad \forall i \quad (6)$$

$$DoD_n \geq target\_value(n) \quad \forall n \quad (7)$$

$$x_{ikl} \geq 0 \quad \forall i, \forall k \text{ and } \forall l$$

$$y_{iju} \geq 0 \quad \forall i, \forall j \text{ and } \forall u$$

Equation (1) i s th e objective fu nction of th e framework. Sal es p rofit, pr oduction cost, tran sportation co st, inventory cost , or dering cost , pu rchasing c ost f or parts and i nventory c ost f or parts are co nsidered in t he proposed objective function. In a special case where a node represents the last-tier supplier, D_{iu} (demand for parts) is zero for all *i* and *u*. On the contrary, d_{kl} (demand for product) is zero for all *k* and *l* if a node represents the final customer. Equation (2) means that the quantity of products made in plants should be larger than the demand of cu stomers. In equation (4), D_{iu} is determ ined according to the decision variable x_{ikl} and the quantity of parts purch ased fro m supplier should be larger than the dem and of parts for plants. Equations (3) and (5) are constraints for production capacity and s upplier capacity, respectively.
In equation (7), DoD_n is the degree of diversification of supply network corresponding to the risk. DoD_1, DoD_2 and DoD_3 are corresponded to the supply disruption risk, the process risk, and the demand risk, respectively. Changes in degree of di versification mean t hat t he changes in *target_value(n)* in e quation (7), and these c hanges lead t he reconfiguration of supply net works. Increased *target_value(n)* means that the e nterprise wa nts more distributed production and procurement plan, and increased value of DoD_n indicates that the enterprise's supply network are more suitable to distribute the risks. In this paper, we used a fuzzy technique to evaluate *DoD*. The detailed descriptions are presented in the next section.

3.2 *DoD* Evaluation Model

DoD_1 is used for multiple suppliers type of configuration (supplier disruption risk), and it is calculated through equation (8).

$$SoC_1 = \sum_{i=1}^{N}\sum_{u=1}^{U}\left\{ f_1(p_1, p_2)\times\left(\sum_{j=1}^{M} y_{iju}\right)\right\} \tag{8}$$

In equation (8), $f_1(p_1, p_2)$ is a fuzzy membership function. Input parameter p_1 is the number of suppliers providing part(u) to plant(i). Input parameter p_2 means the amount of the supplies of parts (u) that are concentrated on a specific supplier. Each input parameter in the fuzzy membership function is calculated as follows.

$$p_1 = \sum_{j=1}^{M} sign(y_{iju}) \tag{9}$$

$$p_2 = \frac{\max\left(y_{iju}, j = 1,\ldots,M\right)}{\sum_{j=1}^{M} y_{iju}} \tag{10}$$

Increment of input parameter p_1 has a positive effect on value of DoD_1, and on the contrary, increment of input parameter p_2 has a negative effect. In equation (10), a denominator means the total quantity of parts (part(u)) purchased by plant(i) from all suppliers in the networks, and a numerator means the maximum value among the quantities of parts(u) purchased by plant(i) from each supplier. The maximum and minimum value of the parameter p_1 is equal to M (the total number of suppliers in the networks) and zero, respectively. The maximum and minimum value of the parameter p_2 is equal to one and $1/M$, respectively. The maximum and minimum values of parameters are important in designing the fuzzy sets.

Although very important, the complicated methods of designing fuzzy sets and rules are beyond the scope of this research. In this paper, thus, very simple fuzzy sets and rules are used. Figure 3 illustrates the examples of fuzzy sets. Fuzzy rules related with input parameters p_1 and p_2 are designed using IF-THEN rules. Based on the fuzzy sets and fuzzy rule, $f_1(p_1, p_2)$ is calculated through the fuzzy inference system and the defuzzification method. See Mun et al. (2009) for a detailed description of fuzzy inference system and defuzzification.

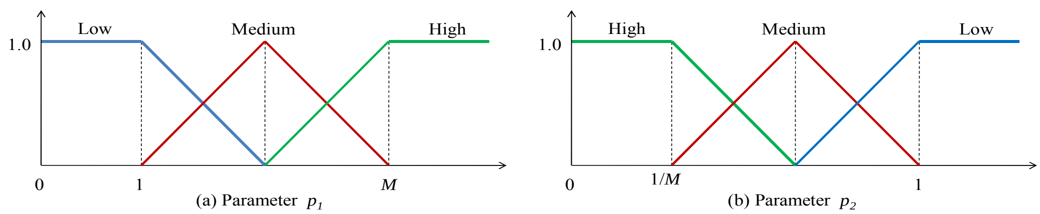


Figure 2. The example of fuzzy sets

DoD_2 is calculated by using equation (11) to (13).

$$p_3 = \sum_{i=1}^{N} sign(x_{ikl}) \tag{11}$$

$$p_4 = \frac{\max\left(x_{ikl}, i = 1,\ldots,N\right)}{\sum_{i=1}^{N} x_{ikl}} \tag{12}$$

$$SoC_2 = \sum_{k=1}^{K}\sum_{l=1}^{L}\left\{ f_2(p_3, p_4)\times\left(\sum_{i=1}^{N} x_{ikl}\right)\right\} \quad (13)$$

DoD_2 has the same structure of fuzzy membership function as SoC_1. Input parameter p_3 is the number of plants providing product(l) to customer(k). Input parameter p_4 means how many productions of product(l) are concentrated on a specific plant. In equation (13), a denominator means the total quantity of product(l) purchased by customer(k) from all plants in the networks, and a numerator means the maximum value among the quantities of product(l) purchased by customer(k) from each plant. The maximum and minimum value of the parameter p_1 is equal to N (the total number of plants in the networks) and zero, respectively. The maximum and minimum value of the parameter p_2 is equal to one and $1/N$, respectively.
DoD_3 is calculated using equation (14) to (16).

$$p_5 = \sum_{l=1}^{L} sign(\sum_{k=1}^{K} x_{ikl}) \quad (14)$$

$$p_6 = \frac{\max\left(\sum_{k=1}^{K} x_{ikl}, l = 1,\ldots, L\right)}{\sum_{k=1}^{K}\sum_{l=1}^{L} x_{ikl}} \quad (15)$$

$$SoC_3 = \sum_{i=1}^{N}\left\{ f_3(p_5, p_6)\times\left(\sum_{k=1}^{K}\sum_{l=1}^{L} x_{ikl}\right)\right\} \quad (16)$$

Input parameter p_5 is the variety of products made in plant(i). Input parameter p_6 means the number of productions that are concentrating on a specific product in plant(i). In equation (16), the denominator means the gross production of plant(i), and the numerator means the maximum value among the quantities of each product produced in plant(i). The maximum and minimum value of the parameter p_5 is equal to L (the number of products in the networks) and zero, respectively. The maximum and minimum value of the parameter p_6 is equal to one and $1/L$, respectively.

4. NUMERICAL EXAMPLE

In this section, to illustrate the proposed framework, a simple numerical example which focuses on the multiple suppliers type (supplier disruption risk) is developed. In the example, a supply network which consists of one customer, two plants (plant(i), i=1, 2), six suppliers (supplier(j), j=1,…,6) and one alternative supplier (supplier(j), j=7) is assumed. One type (product(l), l=1) of final product and three types (part(u), u=1, 2) of parts are also assumed. To produce one unit of product(1), one unit of part(1) and two units of parts(2) are needed. The basic information of the example supply network is summarized in Table 5. To simplify the example, the information related with transportation and inventory is omitted in Table 3. Demand of product(1) are 300 units/month, and unit price of product(1) is $6.

The production and procurement plan of the example supply network without considering the diversification is summarized in Table 4. Values in Table 6 refer to the quantities of part(u) purchased by plant(i) from supplier(j). They are optimized values for the maximization of profit. According to the production and procurement plan represented in Table 4, the expected profit (maximum profit) of the whole supply network is $239 and the DoD_1 is 147. According to the proposed objective function considered the diversification, the procurement plans are changed as shown in Table 5. According to the production and procurement plan represented in Table 5, the expected profit of the whole supply network is $239 and the DoD_1 is 300.

The DoD_1 value of the original supply network is 147, and the reconfigured supply network is 300. In both cases, the expected profit of the entire supply network is the same. As a result of reconfiguration (diversification of production and procurement), the number of suppliers is increased, and the supply of parts(u) is distributed in multiple suppliers maintaining the profit as expected. For example, the supply of part (2) for plant(2) is distributed in supplier(4), supplier(5), supplier(6) and supplier(7) when *target_value(1)* is 300, while the supply of parts(2) for plant(2) is concentrated in supplier(4) before reconfiguration. This means that the reconfigured supply network is more robust to the supply disruption risk than the conventional supply network.

Table 3. Basic information of example supply network

Unit processing cost of plant(*i*); As_{il}	As_{11} = \$ 1.9, As_{21} = \$ 2.0							
Capacity of plant(*i*); c_{il}	200 unit/month (for all *i*)							
Price of part(*u*); C_{iju} (C_{1ju} = C_{2ju})		*j*						
		1	2	3	4	5	6	7
	u 1	\$ 1	\$ 1.1	\$ 1.2	Null	Null	Null	\$ 1.2
	u 2	Null	Null	Null	\$ 1	\$ 1.1	\$ 1.2	\$ 1.2
Capacity of supplier(*j*); C_{ju}		*j*						
		1	2	3	4	5	6	7
	u 1	110	110	110	0	0	0	100
	u 2	0	0	0	220	220	220	100

Table 4. The production and procurement plan before reconfiguration (*target_value(1)* = 0)

		Supplier(*j*)					
		1	***2***	***3***	***4***	***5***	***6***
Plant(*i*)	**Part(*u*)**						
1	***1***	10 110		80	0	0	0
	2	0 0		0	20	220	160
2	***1***	100	0 0		0 0		0
	2	0 0		0	200	0	0

Table 5. The production and procurement plan after reconfiguration (*target_value(1)* = 300)

		Supplier(*j*)						
		1	***2***	***3***	***4***	***5***	***6***	***7***
Plant(*i*)	**Part(*u*)**							
1	***1***	70 103 23			0	0	0	4
	2	0 0 0			99	189	105	11
2	***1***	40	7 0 0			0 0		53
	2	0 0 0			125	31	13	31

If the enterprise wants more flexible procurement, higher *target_value(1)* can be selected. Because the *target_value(1)* has an effect on the expected profit of a supply network as well as the degree of diversification reconfiguration, however, the decision makers should be careful in determining the level of *target_value(1)*. If the enterprise is willing to maintain the current profit level, decision makers can set the *target_value(1)* as 300. On the other hand, he would select the higher *target_value(1)* if the enterprise wants more distributed procurement while willing to bear the decline of the expected profit. The examination of the relationship should be conducted before determining *target_value(1)* because the relationship between *target_value(1)* and expected profit varies according to the characteristics of a supply network.

5. CONCLUSIONS

In this paper, we developed a framework for diversifying the production and procurement. The new framework is composed of a supply network architecture, *DoD* evaluation models, and an objective function. The *DoD* evaluation model was developed using a fuzzy technique and embodied in the objective function. The production and procurement plan of the supply network was optimized to diversify the supply network risks. Planning of production and procurement was demonstrated in a numerical example of the multiple suppliers type, and the result indicates that the proposed framework is effective in distributing the supply network risks.

The main contribution of this research is that the *DoD* corresponding to the risks was evaluated in terms of numerical value, demonstrating that the diversification of supply networks can be evaluated in terms of numerical comparison. For further studies, we are planning to extend the mathematical models to consider more factors related with strategies and risks. In addition, evaluation methodologies for *DoD* need to be developed to support diverse risks that have not been considered here because they are beyond the scope this paper.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0077660) and (2009-0087861). The authors would like to express their gratitude for the support.

5. REFERENCES

Burnes B.and New S. (1996). Understanding supply chain improvement. *European Journal of Purchasing & Supply Management, 2(1):* 21-30.

Carter J.R., Smeltzer L. and Naraslimhan R. (1998). The role of buyer and supplier relationships in integrating TQM through the supply chain. *European Journal of Purchasing & Supply Management, 4(4):* 223-234.

Chopra S. and Sodhi M.S. (2004). Managing risk to avoid supply chain breakdown. *MIT Sloan Management Review, 46(1):* 53-62.

Christopher M. and Lee H. (2004). Mitigating supply chain risk through improved confidence. *International Journal of Physical Distribution and Logistics Management, 34(5):* 388-396.

Demeter K., Gelei A. and Jenei I. (2006). The effect of strategy on supply chain configuration and management practices on the basis of two supply chains in the Hungarian automotive industry. *International Journal of Production Economics, 104(2):* 555-570.

Harland C.M., Lamming R.C., Zheng J. and Johnsen T.E. (2001). A taxonomy of supply networks. *The Journal of Supply Chain Management, A Global Review of Purchasing and Supply:* 21-27.

Kleindorfer P. R. and Saad G.H. (2005). Managing disruption risks in supply chains. *Production and Operations Management, 14 (1):* 53-68.

Mun J., Shin M. and Jung M. (2009). A goal-oriented trust model for virtual organization creation. *Journal of Intelligent Manufacturing,* DOI 10.1007/s10845-009-0293-7.

Oke A. and Gopalakrishnan M. (2009). Managing disruptions in supply chains: A case study of a retail supply chain. *International Journal of Production Economics, 118(1):* 168-174.

Su J.C.P., Chang Y.L. and Ferguson M. (2005). Evaluation of postponement structures to accommodate mass customization. *Journal of Operation Management, 23(3-4):* 305-318.

Tang C. and Tomlin B. (2008). The power of flexibility for mitigating supply chain risks. *International Journal of Production Economics, 116:* 12-27.

Wildemann H. and Carlson J.G. (1987). Implementing Just-In-Time concepts into European companies. *Engineering Costs and Production Economics, 13:* 27-37.

HEADING TOWARDS SUPPLY CHAIN PREDICTABILITY AND SECURITY

Prof. Dr. Frank Arendt[1], Dr. Nils Meyer-Larsen[2], and Rainer Müller[3]

[1]Institute of Shipping Economics and Logistics (ISL)
t.i.m.e.Port II
Barkhausenstraße 2
Bremerhaven, Germany
e-mail: arendt@isl.org

[2]Institute of Shipping Economics and Logistics (ISL)
t.i.m.e.Port II
Barkhausenstraße 2
Bremerhaven, Germany
e-mail: meyer-larsen@isl.org

[3]Institute of Shipping Economics and Logistics (ISL)
t.i.m.e.Port II
Barkhausenstraße 2
Bremerhaven, Germany
e-mail: mueller@isl.org

Abstract: This paper describes how chain visibility of intermodal container chains can be achieved by different measures. Different event sources like Radio Frequency Identification (RFID), Container Security Devices (CSD), and others can enable Supply Chain Event Management (SCEM) in intermodal container transports in order to support operators to meet present and upcoming challenges like handling increasing cargo volumes, the necessity for improved monitoring and stronger security demands. The Institute of Shipping Economics and Logistics (ISL) is currently examining these areas in the course of several research projects.

In the project INTEGRITY the Shared Intermodal Container Information System (SICIS) is developed, allowing authorised companies and authorities to access planning and status information of selected containers and consignments. Proactive planning according to the SCEM approach allows problems to be forecast well before they might occur. Matching logistics data with security information from different sources, together with the integration of the AEO (Authorised Economic Operator) concept, are basic measures.

1. INTRODUCTION

The term Supply Chain Visibility means the intention to close information gaps still existing in international intermodal container chains in order to achieve a better reliability and predictability. Today, the high numbers of involved parties (from industry and administration) are not informed in such a way they can perform their tasks on planning and monitoring in an optimal way. Closing information gaps can be obtained in different ways:

- using classical EDI using EDIFACT or XML interfaces which is already a challenge as such but may still lead to time delays between an event occurred and the respective message
- receiving information in an automated way, e.g. by using identification technologies such as RFID or Automatic Identification System (AIS), positioning technologies such as GPS or Galileo signals, and the related communication technologies either land based, via GSM or satellites
- intelligence for analysing information in order to detect the interesting ones, e.g. comparing planning and actual status information automatically in order to identify the deviations where the planner has to act (this is what we call Supply Chain Event Management)
- performing risk analyses based on these information for security risks (as mainly performed by authorities) as well as logistics risks (e.g. not to be able to deliver the container on time).

This paper will deal with the three last issues.

After the decrease in container throughput due to the economic crisis, container numbers are expected to increase again in the near future. Thus, container terminals and transport operators will again have to deal with several challenges like increasing cargo volumes and security demands that put additional burden on them but offering potentials for process optimisation at the same time. Drivers are:

- Commercial: how to cope with continuous rising cargo volumes to be handled
- Legal/Security: how to deal with new security rules and regulations for fighting against terrorism and the change of responsibilities in the chain
- Technical: how to best integrate technologies such as RFID transponders for container identification and electronic seals combining the benefits of classical bolt seals with RFID capabilities.

The usage of RFID and further event sources can enable Supply Chain Event Management (SCEM) in order to support operators to exploit these challenges.

2. RFID IN CONTAINER TRANSPORT

The drivers in the above mentioned areas are twofold: Security and logistics optimisation. While the first is dominated by regulations mainly initiated by security efforts by the US Government, the latter is facing at logistics optimisations and cost reduction potentials due to automated identification procedures throughout the logistics network. Especially RFID can on the one hand raise the security level by the introduction of electronic seals (eSeals) and on the other hand develop optimisation potentials through the automated container identification by using container tags.

The Institute of Shipping Economics and Logistics (ISL) accommodated the increasing importance and relevance of these topics by the formation of its competence area "AutoID and Security in Container Transport" at the ISL branch at the city of Bremerhaven, Germany. The competence area concentrates on the design of concepts as well as the development of pilot projects for the implementation of automatic identification systems and their connection to business and EDI processes, e.g. for container handling, as well as for the building of surveying systems of intermodal networks for the purpose of logistics and security.

The projects PIER and CHINOS examined and validated the capabilities of RFID and SCEM throughout intermodal container transport chains under the leadership of ISL. Due to the fact that container transport is a global business, of course it is useless to perform tests with proprietary applications. The employed technology must in fact be based on the existing global standardisation efforts. As an example, we expect the container tag to be specified as a passive UHF transponder working in a frequency band from 860 – 960 MHz. The global background is established by the standard ISO 10891. The eSeal, however, is specified by the standard ISO 18185, which is a published international standard.

The project PIER, supported with Bremen subsidies by the Bremerhavener Gesellschaft für Investitionsförderung und Stadtentwicklung mbH (BIS) and Deutsche Telekom, demonstrated the feasibility of the application of RFID technology at the terminal interfaces of North Sea Terminal Bremerhaven (NTB). In the context of the project, several scenarios have been defined and implemented in order to examine the practical application of RFID. The scenarios involved a barge transport from the port of Bremen to the Bremerhaven container terminal, containers transported by truck within the area of northern Germany, and rail operations at the container terminal. As an example, during the rail test, several wagons and containers of an incoming train were equipped with RFID tags as shown in Figure 1. After the train was pushed into the terminal, the tags were scanned by the terminal's rail checker during the regular train checking procedure using a handheld RFID reader.

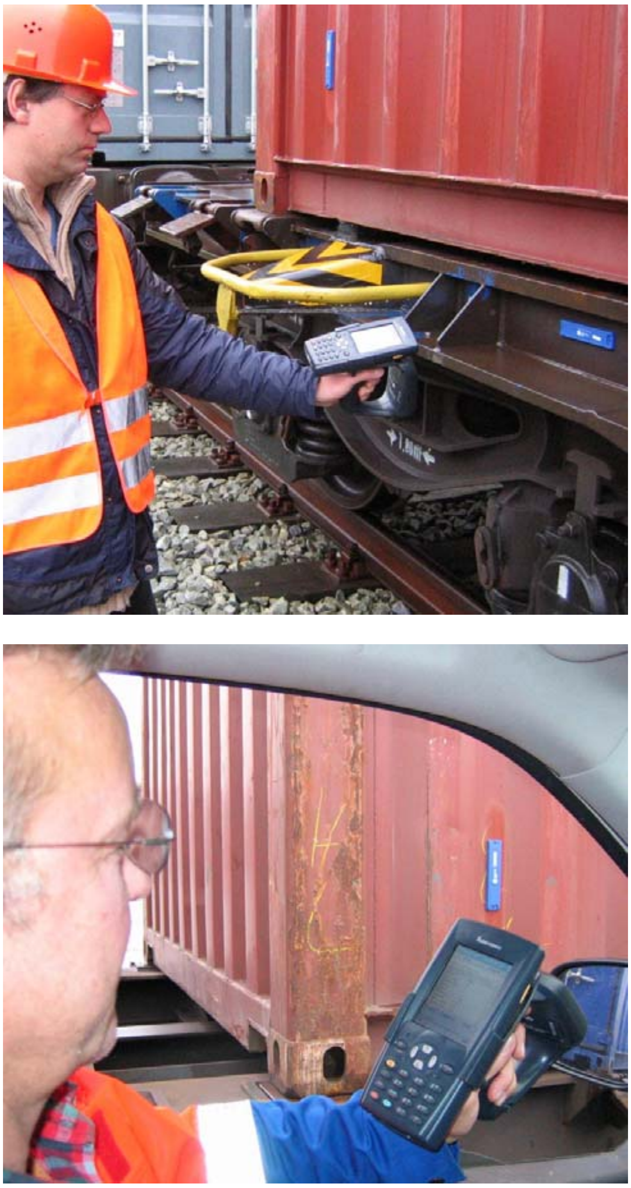


Figure 1. Mounting and readout of container and wagon tags and readout of the tags during the regular train check

Within this project, passive transponders working in the UHF frequency range were used which contained data identifying the containers such as container number and container type. The transponders were read out by handheld RFID reading devices at the terminal's interfaces such as quay, truck gate, and the rail operations facilities. The read data was processed by the software developed within the PIER project and integrated into the terminal's business processes. All tests concluded successfully. The container tags could be read out in all scenarios under all circumstances and the handling of the mobile readers could be integrated smoothly into the existing business processes. These results demonstrate that the new technology leads to considerable optimisations of the processes regarding the acquisition of container data.

Based on the results of the PIER project, CHINOS pursued a much broader strategy. CHINOS stands for "Container Handling in Intermodal Nodes - Optimal and Secure!", was co-funded by the European Commission in their 6th Framework Programme and ran from 2006 until spring 2009. Thirteen partners from Germany, Greece, and Austria, amongst others the dbh Logistics IT AG (Bremen), Eurogate Technical Services (Bremerhaven), i2dm Consulting and Development (Bremen), Tricon Consulting (Traun /Austria), T-Systems (Karlsruhe), and the National Technical University of Athens, examined the application of RFID technology and organisational issues as well as the integration of RFID into existing processes and systems throughout the complete container supply chain. Validation partners were NTB, the rail operator POLZUG, the shipping company Team Lines, Cargo Center Graz from Austria and the port of Thessaloniki together with GAC Shipping from Greece.

As explained above, there were several drivers for this project: on commercial, legal and technical level. All these drivers formed a complex area that the CHINOS project supported:

2.1 Commercial drivers

Before the economic crisis cargo handling figures in container ports were constantly rising. The "Top 50" container ports (representing about 70% of the World container traffic) increased their handling volume in 2003 by 13.7 per cent compared to 2002. Also European ports participated in this growth. Studies predicted an actual average growth of 7% p.a. so that some terminals had to cope with doubling the number of handled containers in seven years.

In order to manage this increase efficiently there are the measures of building new infrastructure, optimising the storage capacity or optimising existing manual procedures, at truck gates, rail gates and in vessel handling. CHINOS has focussed to support the last one by automating procedures.

2.2 Legal and security drivers

After 9/11, several regulations and laws have been implemented to enhance the security in goods transport. Most important one is the introduction of the ISPS Code in port terminals and onboard vessels. Terminals had invested huge amounts of money to become ISPS compliant. But the ISPS code is not the end of the story. In the future, also inland terminals (road/rail or road/inland waterways) will be affected.

Additionally, the replacement of bolt seals with electronic seals (at least in the US trade) is discussed on various political and industrial levels. If becoming mandatory, new requirements appear for the terminal operators.

2.3 Technical drivers

In container transport, several new technologies are in discussion, which create challenges for terminal and transport operators. The RFID technology for containers is currently being discussed in two areas: for identifying the container („container tag" or "license plate") as well as to check the intactness of the seal ("electronic seal") – both to be read electronically and contactless. Electronic licence plates (ISO 10891) as well as electronic seals (ISO 18185) were brought into the standardisation bodies at ISO. For the application of these technologies, big ocean liners (being also the majority of the container owners) will be the forerunners; feeder service operators, terminals, inland transport operators the followers.

For hinterland transports the transport means should be equipped as well. In rail transport, the full benefit is obvious if containers and wagons are equipped with transponders so that the assignment between these two entities can be automatically recorded without the necessity to walk along the train and perform this task manually.

Processes can be optimised and accelerated tremendously by using automatic identification and condition checks with contact free reading possibilities (container RFID tags, electronic seals, optical checks) without requiring human intervention. This combination of commercial and security issues in one approach made CHINOS quite unique. CHINOS terminal operators are able to optimise their storage space and to enhance the integration of transport modes along intermodal logistics chains by redesigning the procedures at their interfaces.

Since the full benefit from new technologies can be exploited only if the total integration of (reengineered) business processes and IT systems will be achieved, CHINOS put a special focus on this integration work and the validation at several European locations, i.e. in the ports of Bremerhaven, Thessaloniki, the Graz freight village and a railway station of the Polzug network.

CHINOS results are ready-to-the-market IT tools as well as technical and organisational recommendations how to efficiently exploit these new technologies to be prepared for the actual and upcoming challenges.

Figure 2. RFID rail tests including container tags and eSeals at Cargo Center Graz

During the first year of the project, the business processes of the involved operators were examined, the validation scenarios were defined, and the technical fundamentals were analysed together with the status of the ISO standardisation process. Furthermore, the different systems to be developed were specified based on the findings from the first project phases and the development phase were pre-pared. After the development phase the CHINOS system has been installed at the validation partners' premises in order to validate under real life conditions by practical tests at all validation locations throughout Europe. Figure 2 shows RFID tests at Cargo Center Graz where the readout of container tags and eSeals was examined using a container train.

2.4 Applicability

The projects showed that an implementation of RFIDs as container tags and electronic seals on a broad scale is possible from the technical side. However, the industry has to invest in the new technology. Main problem here seems that the costs and benefits are not balanced among the partners. Shipping companies have to invest in equipping all their containers with tags and using electronic seals, the terminals have to invest in readers (of course not before a significant number of containers has been equipped with RFIDs) and can automate their processes. Still there is no strategy of the world leaders for the uptake of this technology.

3. SUPPLY CHAIN EVENT MANAGEMENT

Another approach of the CHINOS project was related to an improved monitoring of the container transport network. A Supply Chain Event Management (SCEM) approach can assist the operators monitoring the transport. To achieve this goal, the physical progress of the transport will be compared with the planned procedure by examining so-called events which occur during the transport and could be generated e.g. using RFID. On the one hand there are expected events such as loading and unloading messages. These and the sequence of their occurrence are clearly defined and can easily be monitored. On the other hand there are events which occur unexpectedly and which may allude to problems such as delay messages or a notice indicating a technical defect. If the transport proceeds according to the original plans, the system will stay passive in order to prevent the user from receiving useless “OK” messages. Only if the system detects a deviation of the physical transport from the planned procedure, the operator will be informed pro-actively which enables him to intervene at an early stage of the problem.

The SCEM system compares the expected events with the actual events and decides on appropriate actions, e.g. in-form the user, here the manager of the intermodal transport chain, in case of problems. The chain manager is enabled to react in time on exceptions. Problems can be coped with soon after their occurrence and before they cause a severe impact to the transport process. Thus, an optimisation of the transport chains will become feasible.

Status events can be obtained via RFID, EDI (e.g. discharge messages from container terminals), via mobile devices (e.g. handheld computers using GPRS communication) or using web interfaces.

ISL has developed an SCEM system for logistic purposes to evaluate how SCEM concepts can be used for intermodal container transports covering several sub-transports performed by different transport modes. For each event which can occur during transport there are decision rules which examine its occurrence on time, delay or total absence.

Depending on the result of these examinations, the SCEM system is able to initiate appropriate actions in a flexible way. It can send emails or SMS which notify their receivers about the occurrence of a specific event. In addition, the user's application system can be triggered such that containers originally associated to a cancelled voyage are marked so they can easily be re-scheduled to another voyage; details can be found in Blecker *et al*., 2007.

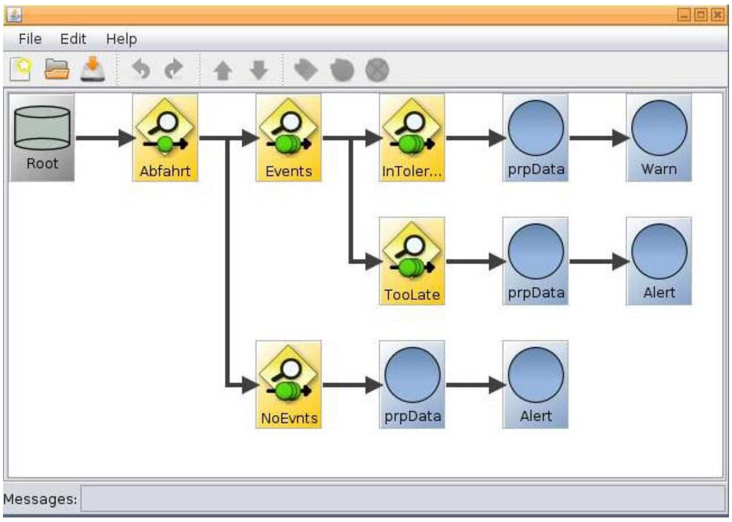


Figure 3. Modelling decision rules for events

Applying the SCEM concept seems to be very attractive in several areas of transport, but is dependent on close to real-time and highly reliable events and status messages.

4. SCEM ENABLES SUPPLY CHAIN VISIBILTY

The visibility of container transports is increased by SCEM software platforms leading to a better reliability and predictability of the transport chain performance.

In the INTEGRITY project, such a platform has been developed which evaluates information from various sources, like RFID for container identification, e-seals or container security devices (CSDs), X-ray inspection and radiation portals to identify illegal contents, satellite tracking of vessels and other vehicles, and external databases for tracing.

4.1 Improving visibility and security

The INTEGRITY project has a wide approach. INTEGRITY is the acronym of "Intermodal Global Door-to-door Container Supply Chain Visibility" being co-funded by the European Commission in their 7th Framework Programme running until spring 2011. It aims at creating Supply Chain Visibility by providing a basis for commercial optimisation and securing of intermodal container chains. Major "clients" of the approach are the commercial participants in the chain (3PLs, cargo owners, exporters, transport and port operators) and authorities (mainly Customs) creating a win-win situation for both of these groups.

Different measures, such as the introduction of the ISPS code in 2004 and the C-TPAT programme in the US, enhanced the security in parts of the international intermodal chain, but a worldwide approach covering the chain from origin to destination is still missing. However, first attempts have been made by the US with the programmes OSC (Operation Safe Commerce) and SST (Smart and Secure Trade Lanes). An important step towards secure operators is the EU Customs Code issued by the Directorate-General Taxation and Customs Union (DG TAXUD) with its AEO (Authorised Economic Operator) approach. Cooperation between Customs Authorities is actually being discussed e.g. in the SSTL project between EU and China Customs Authorities which is closely linked to INTEGRITY. Here, issues of the Customs-to-Customs cooperation will be tackled also from the industry's perspective supporting Customs-to-business and business-to-business cooperation.

If Customs Authorities agree on a mutual recognition of their procedures and a common set of data facilitating pre-arrival clearance before the cargo arrives at its destination, this will speed up the whole process and – even more important – lead to an improved reliability and predictability of the whole chain. Due to the active involvement of Customs services along the demonstration chains from China to Europe as well as the close link to the EU/China Customs Project, the ease of administration together with supporting measures and incentives, e.g. the green lane for supervised secure transports, is covered as well.

The expected benefits are significant: door-to-door chains will become more secure and smooth. All target groups will be satisfied in one approach. Specific analyses on the benefits for all players in the chain analysing actual bottlenecks and performing before-after comparisons including the related costs for such a service are part of the project.

INTEGRITY is an integration project. Although a lot of building blocks are existing, most of the above mentioned technologies have been run through technical feasibility tests without tackling the integration into a common concept on the level of business processes, legal and administrative changes and possible incentives when using them in a consistent and reliable manner. The combination of existing technologies and new business processes together with legal and administrative agreements between administration/Customs and industry/logistics create a win-win situation for both target groups.

4.2 Enhancing security by using technical measures

Different organisational and technical measures can enhance the security of the chain and support the Supply Chain Visibility – also for logistics purposes - at the same time. These are:

- working with trusted parties (AEOs, authorized economic operators)
- using auto-ID methods like RFID for containers
- using X-ray inspection or container imaging (content) facilities
- using radiation portals to identify nuclear materials
- using e-seals or container security devices monitoring the door condition, light, temperature, humidity, radiation, chemicals, etc.
- using satellite tracking of vessels and other vehicles
- using databases with tracing and event information and intelligent algorithms to detect possible risks
- using EDI or web services to perform validity checks with external databases (e.g. owned by transport operators).

In addition to these measures, relevant business processes, e.g. for gaining permissions from authorities, are analysed and adapted for exploiting the technology potentials in an optimal way.

The full-scale integration of IT systems along the chain enables the creation of the so-called Shared Intermodal Container Information System (SICIS) containing either the data itself or links to the data providers (such as port community systems, shipping lines, port authorities) allowing fast and reliable access to the planning data and status information of selected transports. Furthermore, SICIS pro-actively informs the relevant user if possible risks were detected during the transport process. An important issue is the careful handling of the data which is consolidated in the described manner for the first time. The design of SICIS allows its connection to any other system like port community systems or existing legacy systems, hence data security is ensured at any time. Therefore, the data owner has the full control of granting access concerning information on "his" transport to the partners in the chain. Templates enable an easy set-up of recurrent cargo flows.

Figure 4. The Shared Intermodal Container Information System (SICIS) will process data from different data sources and communicate with various platforms

The following scenario will clarify the INTEGRITY approach: A container is stuffed and sealed at a Chinese factory by an AEO certified partner. Another AEO transports the container to the ISPS (International Ship and Port Security Code) certified port of Yantian. Inside the terminal, the seal is checked, and the container is examined with respect to radioactive contents and potentially scanned. After the container was loaded on a vessel operated by an ISPS certified shipping company, all relevant information about the transport including the inspection results is forwarded to the Customs of the importing country. On that basis, Customs can decide on the necessity of a physical inspection of the container already during the sea transport, which will lead to a possible pre-arrival clearance. This will speed up the import process and reduce the risk of delays caused by Customs inspections. As a consequence, the on-carriage can be planned with a much higher accuracy and predictability.

As a first step, SICIS has been implemented at a trade lane starting in mainland China, using the ports of Yantian and Hong Kong on the Chinese side and the ports of Rotterdam and Felixstowe on the European side until the final destination or warehouse. Nevertheless, it is possible to transfer the respective experiences and test results to any other corridor worldwide.

During the INTEGRITY project, tests are carried out using a large number of containers. Rather than focusing on a big bang implementation, INTEGRITY chose a stepwise approach:

- Demonstration phase 1 focused on container data from port networks started in September 2009
- Phase 2 integrated AIS data from vessels the tracked containers were loaded (cp. Figure 5)

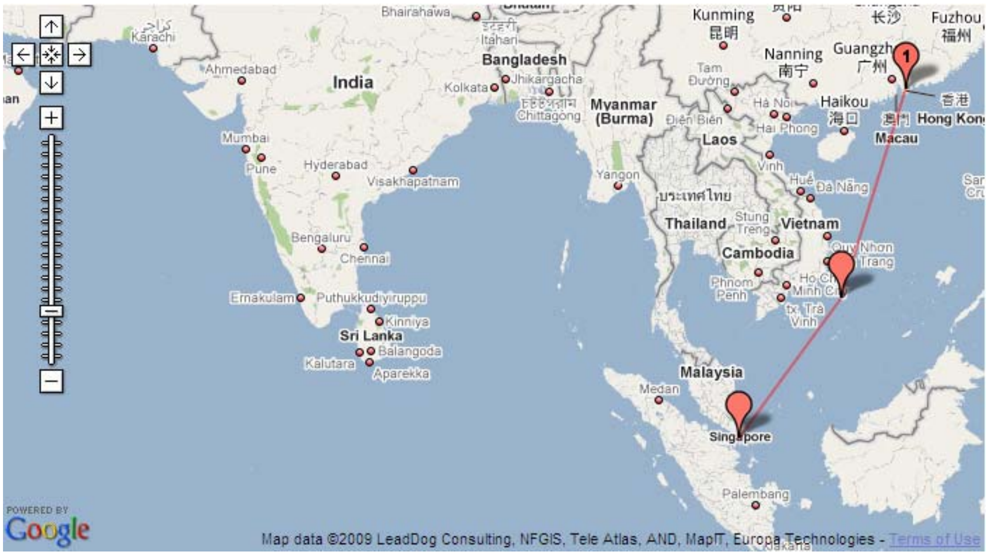


Figure 5. Visualization of vessel events using standardised tools

- Phase 3 added container status and container position data from Container Security Devices (cp. Figure 6)

Figure 6. The first CSD equipped container in preparation for its trip to Europe

- Phase 4 (the actual one) enhanced the system by consignment data.

Recently, several cooperation agreements have been signed in order to extend the INTEGRITY scope to other monitoring technologies, trade lanes and terminal operators.

Although a high data quality during the voyage can be obtained using CSDs which can hardly be achieved by other means, INTEGRITY is not dependent on using that technology. The reason is that it cannot be expected that a significant number of containers will be equipped with CSDs in the near future. Hence, as fallback solution covering also those containers without CSDs, a manual trigger event of the voyage start and end (via web interface) is feasible.

SICIS is achievable to large and SME players in the container business – these are represented by the project partners DHL and BAP Group Ltd. It is the clear intention of the project partners to transfer the pilot system into a real business tool.

5. CONCLUSION

Supply Chain Visibility in intermodal container chains is important for logistics and security control reasons. The paper has shown how technologies (RFIDs, CSDs, AIS) and methodologies (SCEM, mutual recognition of processes) can contribute to this approach. However, a global take-up faces the challenges of investments and process changes.

Innovative technologies like Radio Frequency Identification (RFID) and Supply Chain Event Management (SCEM) can support operators to exploit present and upcoming challenges like handling increasing cargo volumes, the necessity for improved monitoring and stronger security demands. The authors suggest the application of RFID technology, especially the container tag and the electronic seal, in order to improve logistics processes and the security level throughout the container transport chain. Several projects proved the feasibility of these technologies.

Furthermore, it was presented how SCEM software platforms can increase the visibility of intermodal global supply chains in order to improve the reliability and predictability of the transport chain.

The results of the still running project INTEGRITY are promising: based on a broad consensus of the major stakeholders the platform SICIS is accepted by industry and Customs partners supporting logistics monitoring and risk assessment tasks. The near future will show if the intention to transfer SICIS into a commercial service will be successful.

6. REFERENCES

APL Logistics (2003) *Adding Security and Value to the Supply chain*, http://www.apl.com/news/documents/security_white_paper.pdf

Aberdeen Group (2006): The Supply Chain Visibility Roadmap – Moving from Vision to True Business Value, Boston 2006

A.T. Kearney (2005): Smart boxes – RFID can Improve Efficiency, Visibility and Security in the Global Supply Chain, Chicago 2005

Blecker, T. et al. (2007): *Key Factors for Successful Logistics*, Erich Schmidt Verlag GmbH&Co., Berlin

Bundesministerium für Verkehr, Bau und Wohnungswesen (Hrsg.) (2005), *Verkehr in Zahlen 2005/2006*, Deutscher Verkehrs-Verlag, Hamburg

Chinos website: http://www.chinos-rfid.eu

Collins, J. (2005) IBM, *Maersk Developing Cargo Tracker*, RFID Journal Sept. 22, http://www.rfidjournal.com/article/articleview/1884/1/1/

Donner, M., Kruk, C. (2009): Supply Chain Security Guide, World Bank, Washington 2009

European Commission (2001), *White Paper — European transport policy for 2010: time to decide*, Office for Official Publications of the European Communities, Luxembourg

INTEGRITY website: www.integrity-supplychain.eu

INTEGRITY/SmartCM Compendium, Deliverable 1.1

International Maritime Organisation (2003), *The International Ship and Port Security Code*, 2003 Edition

MaTIB website: http://www.matib.de

Savi website: http://www.savi.com

Smart & Secure Tradelanes (2003): Phase One Review – Network Visibility: Leveraging Security and Efficiency in Today´s Global Supply Chains, 2003

World Customs Organisation (2007): WCO SAFE Framework of Standards, 2007

Session D2: Design & Operation 3

·Day1: Sep. 15 (Wed.)

·Time: 16:40 - 18:00

·Chair: Gang Chen

·Room: Azalea & Lilac, 5F

INTEGRATED TACTICAL PLANNING PROBLEM IN MARINE CONTAINER TERMINALS

Gang Chen*[1,2] and Zhongzhen Yang[2]

[1]Center for Maritime Research and Innovation
University of Southern Denmark
Niels Bohrs Vej 9,
Esbjerg, Denmark
Email: gch@sam.sdu.dk

[2]Transport Management College
Dalian Maritime University
Linghai Road, No.1
Dalian, China
Email: yangzhongzhen@263.net

Abstract: This paper addresses the tactical planning problem in an intermodal container terminal, which covers berth allocation, storage space allocation and gate management. There are two typical schemes for tactical operation planning: one in a top-down structure and the other in an integrated parallel structure. The former can utilize the quay side resources but has a weakness in coordinating landside operations. We propose a model for integrated tactical planning, and introduce the wrap round effect into it due to the weekly rolling characteristic of terminal operations. A genetic algorithm heuristic is developed to find a near optimal solution. We compare the proposed integrated model with the independent top-down model by running numerical experiments. The result shows the top-down planning could lead to serious congestion when yard capacity or gate capacity is not sufficient, while the integrated planning can achieve a better performance with less waiting time for both ships and trucks.

Key words: container terminal, integrated tactical planning, gate congestion, berth allocation, storage space allocation

1. INTRODUCTION

The container transportation demand has been continuously growing in last decades. As an intermodal interface between sea and land, mega container terminals serve thousands of ships and millions of trucks each year. We can see that seaside demand gets much more emphasis than landside trucks, because seaside demand links directly with the income of a terminal. In research, seaside container movement is also believed to be the major bottleneck of the whole system. However, the neglected landside truck demand has been generating serious congestion in some major seaports, which in turn limits the productivity of a terminal. Therefore, how to balance the service for ships and trucks in order to achieve the overall utilization of a container terminal becomes an important question.

Container terminal operation planning in a tactical level could be broadly divided into three parts, as shown in Figure 1. *Berth allocation problem* (BAP) refers to the problem of allocating ships to berths or quay locations; *storage space allocation* (SSA) refers to the problem of allocating containers to sub-blocks; *gate time window assignment* (GTWA) refers to the problem of assigning time windows for truck arrivals. These parts are in a top-down structure: the output of BAP is the input of SSA, and the output of SSA is the input of GTWA. This top-down structure can utilize the quay side resources, but pays relatively little attention to the land transport. In fact, landside truck congestion has been becoming serious in many seaports. For example, one empirical study documents that more than 40% of the transactions (pick-up and drop-off of the containers) in the ports of Southern California had to wait more than 2 hours (Barber *et al.*, 2001). In China although the terminals provide 24×7 service throughout the year, large numbers of container trucks wait very often in port areas to be served. Landside congestion does in turn limit the capacity of quay side resources, but this negative impact is not considered in the top-down structure. More planning collaboration between the three parts is expected to increase overall terminal utilization, as Stahlbock and Voss (2008) mention ‘improved terminal performance cannot necessarily be obtained by solving isolated problems but by an integration of various operations connected to each other’.

This paper addresses the integrated tactical planning problem in an intermodal container terminal. We recognize two feedbacks from SSA and GTWA to BAP, which should be added into the planning structure. We find it impossible to add them into the top-down structure, because that will make a loop which may run too many times before finding an optimal solution. Therefore we propose an integrated parallel planning model, in which BAP, SSA and GTWA are in the same level. They start to draft plans simultaneously and then exchange preliminary results with each other; based on the exchanged information the plans are completed and finally get confirmed if the infrastructure capacities can afford.

In this paper, the proposed integrated parallel tactical planning is compared with the top-down tactical planning through numerical experiments. We first develop the integrated planning model by taking the terminals in northern China as a prototype, where GTMA adopts the 'shared vessel-dependent time window method'[a] rather than the 'gate appointment system'[b]. For an introduction and analysis to the GTWA solutions, we refer to Giuliano and O'Brien (2007), Huynh (2009) and Chen and Yang (2010). Then we improve the model by introducing the weekly wrap around effect, because most ships arrive on a weekly schedule. We choose a service-oriented performance indicator[c], ship turn-around time, as the model objective. At last, we develop a genetic algorithm heuristic in order to find a near optimal solution.

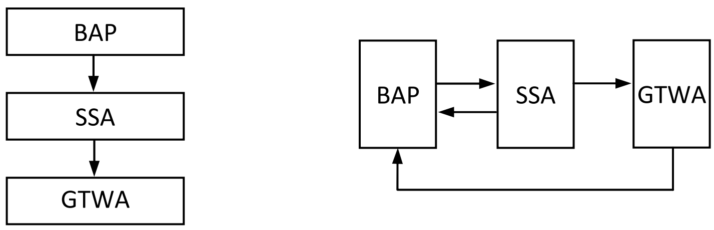


Figure.1 Top-down Structure and Integrated Parallel Structure of Tactical Planning in Terminals

2. LITERATURE REVIEW

There is a huge literature in the area of marine container terminal modelling to date. For an excellent review on port related operations research, we refer to the research work done by Stahlbock and Voss (2008) and Steenken *et al.* (2004), which are based on compiling over 300 literature sources. We can see that a few studies focusing on the overall terminal performance from an integrative view. For example, Murty *et al.* (2005) describe the daily operations of a container terminal in details and discuss the mathematical models and algorithms used in designing a decision support system, which is used to support the daily operations of terminals in Hongkong. Alessandri *et al.* (2007) adopt a system of queues to represent the standing and movements of containers inside intermodal terminals, and propose a linear discrete-time model in order to optimize the container transfer operations. These integrative researches mostly focus on daily operations of terminals rather than the tactical planning problems.

An increasing number of tactical planning researches on separated problems in terminals appear in the last several years. Tactical planning deals with midterm planning issues and afterwards is a primary diver and an input of the operational planning phase. Moorthy and Teo (2006) address a continuous BAP problem, berth template design, which arises in transhipment container terminal. The problem deals with the allocation of favourite berthing locations (home berths) to vessels that call at the terminal based on a weekly schedule. The authors propose two procedures to build good and robust templates and evaluate them by numerical simulations. Cordeau *et al.* (2007) address the service allocation problem, a tactical problem arising in the yard management of a container transhipment terminal. They define a service as the sequence of ports visited by a vessel, where shipping companies usually ask a dedicated specific yard area and a specific berth from a terminal. The objective is to minimize the volume of container rehandling operations inside terminal yards by optimizing the home berth for each service. Giallombardo *et al.* (2010) integrate the discrete BAP problem and the quay crane assignment problem in transhipment container terminals at a tactical level. The objective is the maximization of quay crane utilization and the minimization of the container housekeeping costs. To solve the problem they develop a heuristic algorithm which combines tabu search methods and mathematical

[a] The vessel-dependent time window method used in northern China, manages truck arrivals with sharable and relatively long gate time windows. Before a vessel arrival, a terminal operator assigns a gate time window for all the related container drop-offs. The truckers just need to deliver their containers in the time window without any preannouncement of their arrivals. After the vessel's departure, another time window is assigned to all the container pick-ups.

[b] Gate appointment system is adopted by some terminals in North America and Europe. Through a web-based information system, a terminal operator announces opening hours and entry quota within each hour, and truckers choose entry times they need.

c There are two classes of container terminal performance measures. One is service-oriented indicators, which measure the customer service levels provided by a terminal, for example ship turn-around time, vessels berthed on time, truck waiting time at gates, etc. the other is productivity-oriented indicators, which measure the productivity of a terminal, such as throughput volume, crane utilization, berth utilization, storage productivity and gate processing rate.

programming techniques. From above researches, we could see two versions of tactical BAP problem: discrete case and continuous case. In this paper we consider the discrete BAP.

Some other related work is about BAP and GTWA problems at an operational level. The operational BAP problem has received a larger attention than the tactical one so far. For discrete operational BAP, we refer to Imai et al. (2001, 2003) and Cordeau *et al.* (2005); for continuous operational BAP, we refer to Imai et al. (2005) and Guan and Cheung, (2004). Regarding GTWA, there are two typical solutions as above mentioned: vessel-dependent time window methods and gate appointment system. This paper considers the former one. Yang *et al.* (forthcoming, 2010) find that truck arrivals for container drop-offs within a time window basically followed a beta distribution, which enables us to predict truck arrivals based on gate time window assignment. Chen and Yang (2010) propose an idea of optimizing time windows as a means of managing truck arrivals in container terminals and develop a heuristic algorithm for finding a near optimal time window assignment, and the objective is the minimization of total operation cost.

3. PROBLEM FORMULATION AND MATHEMATICAL MODELS

In this section, we first formulate the top-down tactical planning problem and then the integrated parallel planning problem. The objective is the minimization of ship turn-around time, which includes waiting time for the berth availability, handling time and extra waiting time due to landside congestion. Since these indicators are not dependent on import container pick-ups, we don't consider import operations in GTWA part. In our models, we make a number of assumptions for simplicity:

(1) Each berth can service one ship at a time and that there are no physical or technical restrictions such as a relationship between ship draft and water depth;
(2) The handling time of a ship depends on the distance between the ship and container storage location, the handling volume and the efficiency of the assigned berth;
(3) Once a vessel is moored, it will remain in its location until all required container processing is done. In practice, any interruption of the container processing during mooring is costly. Thus, this assumption is valid in practice;
(4) Both terminals and trucking companies provide 24×7 service through the whole year, which is true in China.

3.1 Top-Down Planning Model

In this section we present a formulation for the top-down tactical planning. All the input data for the model are deterministic, as follows:

I	set of ships
J	set of berths
K	set of yard blocks
P	set of time steps in the planning horizon (24×7 hours for a weekly template)
A_i	arrival time of ship i
V_i	handling volume of ship i (TEU)
r_i	the percentage of loaded export containers over the total handling volume of ship i
B_k	storage space for export containers in the yard block k (TEU)
G	gate processing rate (trucks/hour)
N_j	handling efficiency of berth j (TEU/hour)
β	truck modal split for the container deliveries;
f	average loading factor of a truck (TEU)

The model also uses the following variables and sets:

- $x_{ij} \in \{0,1\}\ \forall i \in I, \forall j \in J$, set to 1 if ship i is assigned to berth j, and 0 otherwise;
- $y_{ik} \in \{0,1\}\ \forall i \in I, \forall k \in K$, set to 1 if the containers of ship i are stored in block k, and 0 otherwise;
- $s_{mi}^{j} \in \{0,1\}\ \forall (i,m) \in I, \forall j \in J$, set to 1 if ship i is scheduled after ship m at berth j, and 0 otherwise;
- $T_i^C \geq 0\ \forall i \in I,$ representing the handling completion time and the departure time of ship i;
- $T_i^S \geq 0\ \forall i \in I,$ representing the time when the terminal begins to accept container drop-offs for ship i;
- $T_i^{ES} \geq 0\ \forall i \in I,$ representing the earliest possible time point of T_i^S;
- $T_i^{LS} \geq 0\ \forall i \in I,$ representing the latest possible time point of T_i^S;
- $T_i^E \geq 0\ \forall i \in I,$ representing the time when the terminal stops to accept container drop-offs for ship i;
- $z_i^w \geq 0\ \forall i \in I,$ representing the waiting time of ship i for the availability of the assigned berth;

- $z_i^h \geq 0 \; \forall i \in I$, representing the actual handling time of ship i at the assigned berth;
- $h_{ij} \geq 0 \; \forall i \in I, \forall j \in J$, representing the expected handling time of ship i if it is moored at berth j;
- $e_{ti} \in \{0,1\} \; \forall i \in I, \forall t \in P$, set to 1 if time step t is within the period $[T_i^S, T_i^C]$, and 0 otherwise;
- $D_t \geq 0 \; \forall t \in P$, representing the number of truck arrivals for export container drop-offs during time step t;
- $q_t \geq 0 \; \forall t \in P$, representing the queue length of trucks at the terminal gate during time step t;
- $w_t \geq 0 \; \forall t \in P$, representing the gate waiting time of the trucks coming during time step t.

3.1.1 BAP Submodel

$$min \sum_{i \in I} (z_i^w + z_i^h) \quad (1)$$

$$z_i^w = max\left(0, \sum_{m \in I} \sum_{j \in J} x_{ij} s_{mi}^j \, T_m^C - A_i\right) \quad (2)$$

$$h_{ij} = 2 + V_i / N_j \quad (3)$$

$$z_i^h = \sum_{j \in J} x_{ij} h_{ij} \quad (4)$$

$$T_i^C = A_i + z_i^w + z_i^h \quad (5)$$

Subject to

$$\sum_{i \in I} x_{ij} z_i^h \leq 24 \times 7 \quad \forall j \in J \quad (6)$$

$$\sum_{j \in J} x_{ij} = 1 \quad \forall i \in I \quad (7)$$

$$\sum_{m \in I} s_{mi}^j \leq 1 \quad \forall j \in J, \forall i \in I \quad (8)$$

$$\sum_{m \in I} \sum_{i \in I} s_{mi}^j = \sum_{i \in I} x_{ij} - 1 \quad \forall j \in J \quad (9)$$

$$x_{ij} \in \{0,1\} \; \forall i \in I, \forall j \in J \quad (10)$$

$$s_{mi}^j \in \{0,1\} \; \forall i \in I, \forall m \in I, \forall j \in J \quad (11)$$

The decision variables of BAP submodel are x_{ij} and s_{mi}^j. The objective (1) minimizes waiting and handling time for every ship. Equation set (2) calculates the waiting time of each ship for the availability of the assigned berth. Equation sets (3) and (4) calculate the actual handling time of each ship at the assigned berth assuming that the related containers are stored in the closest yard block. Equation set (5) calculates the expected completion time of each ship's handling, regardless of the influence of yard operations and gate operations. Constraint (6) ensures that the workload (hours) of each berth is not over the planning horizon. Constraint (7) ensures every ship must be served at some berth. Constraint (8) and (9) ensure that every ship is scheduled following another ship at the same berth, except only one ship that is scheduled as the first one.

3.1.2 SSA Submodel

$$max \sum_{i \in I} (T_i^E - T_i^{ES}) V_i - \lambda \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} x_{ij} y_{ik} |j - k| \quad (12)$$

$$T_i^E = A_i + z_i^w \quad (13)$$

$$T_i^{LS} = T_i^E - max(6 + V_i r_i / G) \quad (14)$$

Subject to

$$T_i^{ES} \leq T_i^{LS} \quad \forall i \in I \quad (15)$$

$$\sum_{i \in I} y_{ik} V_i r_i e_{ti} \leq B_k \quad \forall k \in K, \; \forall t \in P \quad (16)$$

$$\sum_{k \in K} y_{ik} = 1 \quad \forall i \in I \quad (17)$$

$$y_{ik} \in \{0,1\} \; \forall i \in I, \forall k \in K \quad (18)$$

$$e_{ti} \in \{0,1\} \; \forall i \in I, \forall t \in P \quad (19)$$

The decision variables of SSA submodel are y_{ik} and T_i^{ES}, and the inputs x_{ij} and z_i^w are obtained from BAP submodel. The objective (12) minimizes the distance between the assigned berth and the containers position in yard for all ships, and then maximizes the gate service time for export container drop-offs. The parameter λ is an extreme high value penalty. Equation set (13) calculates the ending point of each gate time window. Equation set (14) calculates the latest possible starting point of each gate time window. Constraint (15) ensures that the possible earliest starting point must be no later than the latest possible starting point of each gate time window. Constraint (16) ensures that, at every time step, the total number of stored export containers does not exceed the storage capacity in a block. Constraint (7) ensures that containers of every ship must be stored in some block.

3.1.2 GTWA Model

$$min \sum_{t\in P} w_t D_t \tag{20}$$

$$D_t = \sum_{i\in I} \int_{t-1}^{t} Beta(1.29, 3.25, [T_i^S, T_i^E])\, V_i r_i \beta / f \tag{21}$$

$$q_t = \max(\, q_{t-1} + D_t - G, 0) \tag{22}$$

$$w_t = \frac{q_t + q_{t-1}}{2} / G \tag{23}$$

Subject to

$$T_i^{ES} \le T_i^S \le T_i^{LS} \;\; \forall i \in I \tag{24}$$

The decision variable of GTWA submodel is T_i^S, and the inputs T_i^E, T_i^{LS} and T_i^{ES} are obtained from SSA submodel. The objective (20) minimizes the waiting time of trucks at terminal gate. Equation set (21) calculates the number of trucks arriving at terminal gate during time step t. Equation set (22) calculates the number of trucks waiting at terminal gate at time step t. Equation set (23) calculates the average waiting time of the trucks arriving during time step t. Constraint (24) ensures that the real starting point must between the earliest and the latest possible starting points of each gate time window.

The major problem of the top-down planning model is that, the impact of gate/yard congestions on BAP is not included. BAP cannot be properly solved without the information about delayed container drop-offs from GTWA and the information about container locations from SSA. These feedbacks are necessary for tactical planning, but as above mentioned, the top-down model will become a loop if we add them in. Therefore, the primary concern of the integrated parallel planning model is to build a complete model without a circular cause and consequence.

3.2 Integrated Parallel Planning Model

In order to figure out why a circular cause and consequence happens, we look closely into an example: when gate congestion (some delayed container drop-offs) causes the related ship waiting, the next ship assigned to the same berth cannot moor, of which the export containers have to stay longer in the yard block; so the yard block cannot serve the new coming containers in time, therefore new gate congestion happens. We can see the major reason is a temporary shortage of storage space, which transmits congestion from one side to another and back again. Therefore we propose a parallel structure for integrated planning model, where BAP, SSA and GTWA exchange information before making a solution; any solution has to go through a storage demand test before getting confirmed.

The inputs and variables are the same of the top-down planning model, plus the following variable instead of h_{ij}:

- $h_{ijk} \ge 0 \; \forall i \in I, \forall j \in J, \forall k \in K$, representing the handling time of ship i at berth j with block k;

The integrated parallel planning model is as follows:

$$min \sum_{i\in I} \left(z_i^w + z_i^h\right) \tag{1}$$

$$z_i^h = \sum_{j\in J} \sum_{k\in K} x_{ij} y_{ik} h_{ijk} \tag{25}$$

$$h_{ijk} = 2 + (1 + \tan\alpha |k - j|)\, V_i / N_j \tag{26}$$

$$z_i^w = max\left(w_{T_i^E}, \sum_{m\in I} \sum_{j\in J} x_{ij} s_{mi}^j\, T_m^C - A_i \right) \tag{27}$$

$$T_i^E = A_i + z_i^w - w_{T_i^E} \tag{28}$$

The above equations consist of the complete integrated planning model with Equation set (5), (14) and (21)-(23). The constraints are the same of the top-down planning model, including Constraint set (6)-(11), (15)-(19) and (24). The objective is same to the BAP submodel, i.e. minimization of total ship turn-around time. The major difference with the top-down planning model is that, extra handling time of ship i due to far container locations in yard is included in Equation set (26) and (27), and extra waiting time of ship i due to gate congestion is also included in Equation set (27), but this extra waiting time will not change the time window assignment, as shown in Equation set (28), otherwise it will probably be a circular cause and consequence again.

4. HEURISTIC ALGORITHM

Solving the tactical planning models with a general-purpose solver is difficult, so a genetic algorithm (GA) based heuristics and some mathematic techniques are developed. In this section, we introduce first a technique to calculate the wrap-around effect and then the GA heuristics for the planning models.

4.1 Wrap-around Effect (WAE)

The weekly wrap-around effect happens in every planning model, for example Figure 2 shows the effect in BAP. We adopt Equation (29) to modify the variable values that go beyond the planning horizon, and use Equation (30) to correct the variable values indexed by actual time. These two modifications are enough for most variables, but not for the conditional variables, as T_i^C and q_t, which seem to bend back on itself in the calculation process. This problem is called 'feedback loop' in control theory, from where we borrow the concepts of 'positive feedback' and 'negative feedback' to make a solution. We find a property of the feedback loop as follows, taking T_i^F as example.

Property 1. When total demand (ship handling time) is not over total supply (quay service hour) within a planning horizon, the calculation of conditional variables (T_i^C) is a negative feedback loop. A negative feedback loop can be solved by running two times in the way that: give a hypothetical minimal value to the first variable (T_1^C) in the set, and run the first loop; then correct the first variable based on the last variable value (T_I^C) obtained from the first loop, and run the second loop to correct all other variables.

$$Variable = \begin{cases} Variable - 24 \times 7 & if\ Variable > 24 \times 7 \\ Variable + 24 \times 7 & if\ Variable \leq 0 \end{cases} \tag{29}$$

$$Variable'_t = Variable_t + Variable_{(t+24\times7)} + Variable_{(t-24\times7)} \tag{30}$$

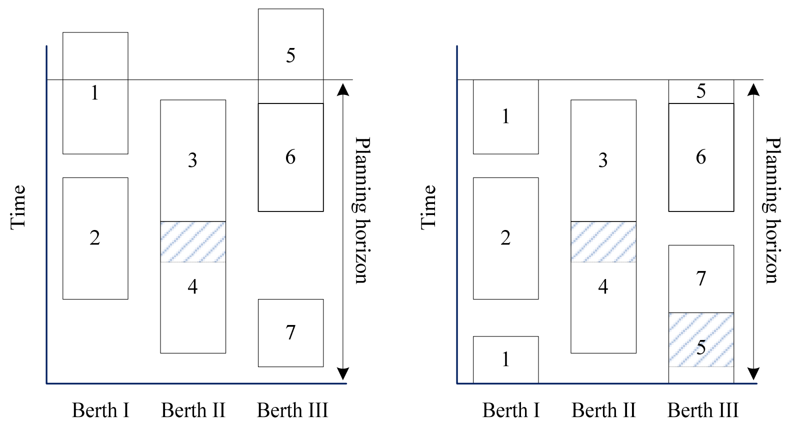


Figure 2. The Wrap Around Effect of BAP Planning

4.2 Heuristics for Top-down Planning Model

We combine GA heuristics with the following mathematic techniques to solve BAP, SSA and GTWA individually.

4.2.1 Assigning s_{mi}^{j} based on x_{ij} (SOX)

SOX is embedded into the heuristic for BAP. Given a set of x_{ij}, s_{mi}^{j} is preliminarily assigned based on 'first come first serve' (FCFS) principle for ships at a specific berth. Since FCFS rule does not promise an optimal solution, we propose two methods to modify the assignment of s_{mi}^{j} for improvement. The first one is shifting ships between berths, as shown in Figure 3. A ship is shifted to one idling berth, if it arrives during the berth idling period and keeps waiting for the availability of another berth. This method is used repeatedly until no ship arrives and waits during any berth idling period. During this process, constraint set (11) is always followed to avoid impossible solution. The second one is swapping two neighbouring ships at a specific berth, as shown in Figure 3. All pairs of neighbouring ships are checked, and swapped by interchanging their service orders if the total ship waiting time at the berth could be shorter.

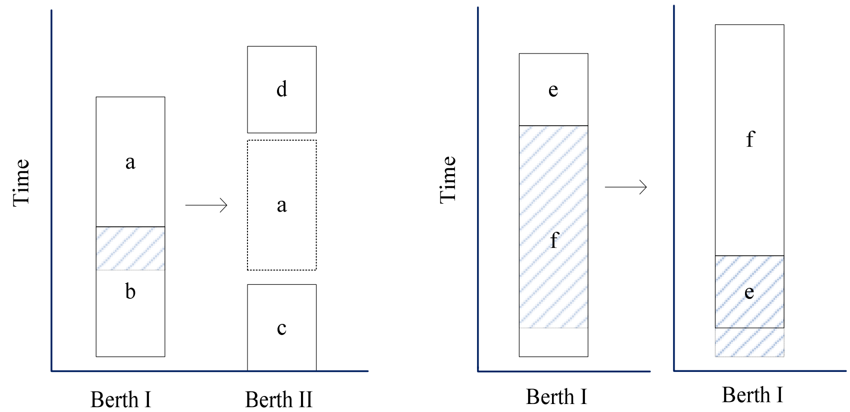


Figure.3 Shift and Swap Operations For BAP

4.2.2 Assigning y_{ik} based on x_{ij} (YOX)

YOX is embedded into the heuristic for SSA. Locating containers far from the ship position will lead to long container transfer distance and long ship handling time. So preliminarily a container location in yard is assumed the same of the berth location for each ship, as shown in Equation (31). This presumed SSA solution should be modified if a block capacity cannot meet its minimal storage demand, where every T_i^S equals toT_i^{LS}. The modification is done in the way that, randomly choose and move one of container locations from the congested block to another uncongested block. This operation is repeated until all yard blocks can meet their minimal storage demands.

$$y_{in} = x_{in} \ \forall i \in I \tag{31}$$

4.2.3 Assigning T_i^{ES} based on y_{ik} (TOY)

TOY is also embedded into the heuristic for SSA. In export container blocks, occupied yard space can only be released by a completion of ship handling. So in a congested block with insufficient space, each T_i^{ES} must be one of the ship handling completions. Figure 4 shows an example, where i ships have their containers stored in this block. $T_i^{ES}$of the example ship cannot be within the diagonal range, which is between $T_i^{LS}$and T_i^C. Since in the diagonal range there are n handling completions, the other $i-n$ points are possible options. Note the first option before T_i^{LS}, in this example T_{n+1}^C, is always feasible, because the remaining storage space at T_{n+1}^C is not less than the one at T_i^{LS}. The earliest feasible option can be identified by checking the remaining storage space in between. At most option $i-n$, the ship's self departure time, could be a feasible option.

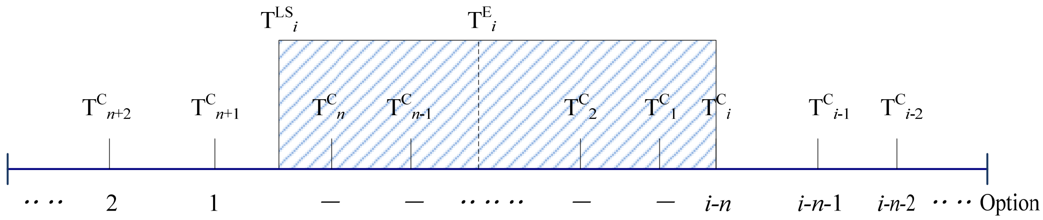


Figure.4 Options for The Earliest Starting Point of A Time Window

4.3 Heuristics for Integrated Planning Model

We propose a GA heuristics combined with the above mathematic techniques to solve the integrated model. The GA heuristic is based on a chromosome representation, which consists of x_{ij} and T_i^{LS} of each ship. The solving steps based on GA are presented as follows.

Outline of the entire solution procedure

Initialization: assign x_{ij} and T_i^{LS}

repeat

1. Solve BAP with process ‘SOX’;
2. Solve SSP via processes ‘YOX’ and ‘TOY’;
3. Calculate the effect of gate congestion via GTWA module;
4. Update x_{ij} and T_i^{LS} with the genetic operations: mutation and crossover;

Until stopping criterion;

Figure. 5 scheme of the heuristic algorithm for integrated planning model

5. Numerical Experiments

We first describe how test instances were generated, and then provide results of the top-down and the integrated parallel models obtained with our heuristics. The heuristics are coded in Matlab version 7.8 and run on Intel(R) 2.10GHz with 1.96GB RAM.

Test instances were generated randomly but systematically. We develop 80 instances with different yard capacities and gate capacities as shown in Table 1, based on one terminal with five berths. A ship data set is generated by an exponential distribution of ship arrivals with an average interval of three hours. The handling volumes of these ships are generated by a uniform distribution with an average value of 1,100 TEU. The distribution of ship arrivals is independent on that of handling volumes. We assume all berths have same handling speed of 100TEU/hour, and export operations and import operations share the yard capacity and the gate capacity by fifty-fifty rule.

Table.1 Parameters for the test instances

Parameter	Value
Number of berths	5
Number of Ships	56
Handling volume [min, max] (TEU)	[10, 2200]
Yard capacity ($\times 10^3$ TEU)	30, 32, 34, …… 48
Gate capacity (Trucks/hour)	200, 210, 220, 230, 240

Table 2 illustrates how solution quality improves by the integrated model. The first column indicates instance size (i.e. capacities of yard and gate), and the next three columns show ship waiting time in BAP, SSA and GTWA models respectively. The fifth column shows total ship turn time in the top-down model solutions. The last two columns of Table 2 show the solutions of the integrated model, one is total ship turn time, and the other is the gap between two solutions. We can see that the top-down model does not have stable solutions, which not allways improves with larger yard capacity. This is because GTWA's objective is not ship turn time, but truck waiting time. Good solutions of GTWA do not necessarily support BAP solutions. This phenomenon indicates that a trade-off between ship turning time and truck waiting time exists when yard and gate capacities are insufficient. Another finding of Table 2 is that containers position in yard is always assigned the same of the corresponding ship's quay position. This indicates that it does not pay off to locate containers far from corresponding ship in a tactical level planning.

Table.2 Comparison of the solutions of the top-down and integrated models

Instance [a]		Top-down				Integrated	
		BAP	SSA	GTWA	total	total	Gap [b]
	200	136	0	−	−	371	−
	210	136	0	−	−	187	−
30 ×	220	136	0	12	156	156	0%
	230	136	0	4	159	147	8%
	240	136	0	0	136	136	0%
	200	136	0	−	−	223	−
	210	136	0	−	−	164	−
34 ×	220	136	0	8	148	144	3%
	230	136	0	−	−	137	−
	240	136	0	0	136	136	0%
	200	136	0	82	263	193	27%
	210	136	0	19	168	151	10%
38 ×	220	136	0	5	139	137	1%
	230	136	0	1	137	136	1%
	240	136	0	0	136	136	0%
	200	136	0	−	−	193	−
	210	136	0	11	185	151	18%
42 ×	220	136	0	12	151	137	9%
	230	136	0	−	−	136	−
	240	136	0	0	136	136	0%
	200	136	0	−	−	193	−
	210	136	0	−	−	151	−
46 ×	220	136	0	1	137	137	0%
	230	136	0	6	163	136	17%
	240	136	0	0	139	136	2%

Instance [a]		Top-down				Integrated	
		BAP	SSA	GTWA	total	total	Gap [b]
	200	136	0	−	−	295	−
	210	136	0	−	−	187	−
32 ×	220	136	0	−	−	156	−
	230	136	0	2	138	138	0%
	240	136	0	0	136	136	0%
	200	136	0	80	318	195	39%
	210	136	0	22	170	151	11%
36 ×	220	136	0	8	150	137	9%
	230	136	0	0	136	136	0%
	240	136	0	0	136	136	0%
	200	136	0	60	199	193	3%
	210	136	0	18	151	151	0%
40 ×	220	136	0	0	136	137	-1%
	230	136	0	0	136	136	0%
	240	136	0	0	136	136	0%
	200	136	0	87	262	193	26%
	210	136	0	32	190	151	21%
44 ×	220	136	0	4	138	137	1%
	230	136	0	0	136	136	0%
	240	136	0	−	−	136	−
	200	136	0	47	187	187	0%
	210	136	0	17	167	167	0%
48 ×	220	136	0	4	139	139	0%
	230	136	0	0	136	136	0%
	240	136	0	0	136	136	0%

[a] Instance = Yard capacity ($\times 10^3$) × Gate capacity.

[b] Gap = (top-down total-integrated total)/ top-down total.

− means that no possible solution is found.

From an overview of Table 2, we could see the integrated planning model really helps to obtain increased terminal performance. The top-down model generates impossible solutions very often, while the integrated model can avoid.

Table 3 shows gate congestion in different instances. Majority data of Table 3 shows that gate congestion is alleviated by larger yard capacity and/or larger gate capacity, except the data of instances 36×200, 38×210and 34×230. These exceptions confirm the trade-off between ship turn time and truck waiting time. On the other hand, this is also one shortage of the integrated model that, saving ship turn time sometimes is obtained by bearing too much gate congestion. For example, comparing with instance 36×200, instance 38×200 generates 3200 hours extra truck waiting time in order to save two hours ship turn time.

Table.3 Gate congestion in the solutions of the integrated model

Gate / Yard($\times 10^3$)	200	210	220	230	240	200	210	220	230	240
	Maximum truck waiting (h)					Average truck waiting (h)				
30	8.1	3.3	2.3	1.5	0.7	2.9	0.9	0.5	0.3	0.2
32	9.9	3.3	2.3	1.5	0.7	2.9	0.9	0.5	0.3	0.2
34	3.0	2.6	1.4	**0.9**	0.7	1.1	0.5	0.3	**0.1**	0.1
36	**1.6**	2.3	1.2	1.0	0.7	**0.7**	0.6	0.2	0.1	0.1
38	2.3	**1.9**	1.2	1.0	0.7	0.9	**0.5**	0.2	0.1	0.1
40	2.3	2.2	1.2	1.0	0.7	0.9	0.6	0.2	0.1	0.1
42	2.3	2.2	1.2	1.0	0.7	0.9	0.6	0.2	0.1	0.1
44	2.3	2.2	1.2	1.0	0.7	0.9	0.6	0.2	0.1	0.1
46	2.3	2.2	1.2	1.0	0.7	0.9	0.6	0.2	0.1	0.1
48	2.3	2.2	1.2	1.0	0.7	0.9	0.6	0.2	0.1	0.1

6. CONCLUSIONS

In maritime container terminals, landside truck congestion has a negative impact on the utilization of quay side resources. But this impact is not considered in the top-down tactical operations planning. Adding it into the top-down structure will make a loop which may run too many times before finding an optimal solution. Therefore we propose an integrated parallel planning model, in which BAP, SSA and GTWA are in the same level. They exchange information before making a solution, so as to avoid loop computation. In the numerical experiments, we can see the integrated planning model really helps to obtain increased terminal performance.

The proposed model can also be used for decision-making problems, for example, how many berths to operate. In this paper the relationship between the number of cranes and ship handling time is not investigated yet, we leave it for next research.

7. REFERENCES:

Alessandr, A., Sacone, S., & Siri, S. (2007). Modelling and Optimal Receding-horizon Control of Maritime Container Terminals . *Journal of Mathematical Modelling and Algorithms* , 109-133.

Barber, D., & Grobar, L. (2001). *Implementing a Statewide Goods Movement Strategy and Performance Measurement of Goods Movement in California.* Long Beach: METRANS Transportation Centre.

Chen, G., & Yang, Z. (2010). Optimizing Time Windows for Managing Arrivals of Export Containers at Chinese Container Terminals. *Maritime Economics & Logistics* , volume 12, 111–126.

Cordeau, J.-F., Gaudioso, M., Laporte, G., & Moccia, L. (2007). The service allocation problem at the Gioia Tauro Maritime Terminal. *European Journal of Operational Research* , 1167–1184.

Giallombardo, G., Moccia, L., Salani, M., & Vacca, I. (2010). Modeling and solving the Tactical Berth Allocation Problem. *Transportation Research Part B* , 232-245.

Giulianoa, G., & O'Brienb, T. (2007). Reducing port-related truck emissions: The terminal gate appointment system at the Ports of Los Angeles and Long Beach. *Transportation Research Part D* , Volume 12, pp.460–473.

Guan, Y., & Cheung, R. K. (2004). The berth allocation problem: models and solution methods . *OR Spectrum* , 75-92.

Huynh, N. N. (2009). Reducing Truck Turn Time at Maritime Container Terminals with Appointment Scheduling. *In Transportation Research Record: Journal of the Transportation Research Board* , No. 3230, pp. 48–56.

Imai, A., Nishimura, E., & Papadimitriou, S. (2003). Berth allocation with service priority. *Transportation Research Part B* , 437-457.

Imai, A., Nishimura, E., & Papadimitriou, S. (2001). The dynamic berth allocation problem for a container port. *Transportation Research Part B* , 401-417.

Imai, A., Sun, X., Nishimura, E., & Papadimitriou, S. (2005). Berth allocation in a container port: using a continuous location space approach. *Transportation Research Part B* , 199-221.

Moorthy, R., & Teo, C.-P. (2006). Berth management in container terminal:the template design problem. *OR Spectrum* , VOL 28; NUMBER 4, pages 495-518 .

Murty, K. G., Liu, J., Wan, Y.-w., & Linn, R. (2005). A decision support system for operations in a container terminal. *Decision Support Systems* , 309– 332.

Stahlbock, R., & Voß, S. (2008). Operations research at container terminals: a literature update. *OR Spectrum* , Volume 30, pp.1–52.

Steenken, D., Voß, S., & Stahlbock, R. (2004). Container terminal operation and operations research – a classification and literature review. *OR Spectrum* , 3-49.

Yang, Z. Z., Chen, G., & Moodie, D. R. (forthcoming). Modelling Road Traffic Demand of Container Consolidation in a Chinese Port Terminal. *accepted by Journal of Transportation Engineering-ASCE* .

Aa. We refer dynamic BAPD simply as BAPD
In case that there are too many ships allocated to one berth i.e. the workload of a berth (total handling time of ships) exceeds the planning horizon P, we reallocate some ships from this berth to another berth randomly. This berth reallocation operation is repeated until workloads of all berths are not higher than the planning horizion.

- $z_i^d \geq 0 \; \forall i \in I,$ representing the extra berthing time of ship i due to the gate congestion;
- There is not a single process for this effect, but it is embed into where it is needed in all processes.
- berth allocation problem
- storage space allocation
- gate processing model

Idea

While the berth and yard templates already attract much attention from researchers, the gate operation template dealing with the landside demand has received less consideration.

The process of template design varies between terminals due to the landside demand management. In the terminals where truck arrivals are not controllable, the gate operation template simply refers to the opening hours, during which the trucks are allowed to come at any time. In this case, the berth template will be drawn first and subsequently used as the input of yard storage planning. But in the other terminals where truck arrivals are under control, either by gate appointment systems[1] or through a time window management method[2], the gate operation template covers the more detailed allocation of the gate capacity, including either opening hours with entry quota or vessel-dependent time windows. By this gate capacity allocation, the pattern of the landside demand will be changed so as to coincide with the supply of terminal operations. Therefore, the template design could be an integrated problem, aiming to minimize the vessels' berthing time, the use of storage space and the external trucks' waiting time.

So, this paper will focus on the template design problem in a congested container terminal adopting a vessel-dependent time windows method for managing the truck arrivals. And the problem formulation is: given a seaside demand, how to make the optimal berth, yard and gate operation templates in order to provide high level service to both the vessels and the trucks.

The objective is to minimize the total cost of the vessels' berthing time, the cargo storage and the external trucks' time.

The berth template design: to decide when and where the vessels to berth, while the berthing time is assumed to depend on the loading and unloading container volume. The below picture is an example of berth solution. The optimization of this part will be a little difficult.

where each part focuses on its own task independently.
However, sometimes the output of lower parts can also influence the upper parts. For example, in the middle part 'storage space allocation', it is possible to allocate the containers far from the ship position due to the short capacity of the closest block. The long distance of container transfer will lead to longer ship handling time than it is supposed to be in the 'berth allocation' part. Another example is that high gate congestion could delay ship operation and departure, thereafter lead to a longer occupation of the yard space.

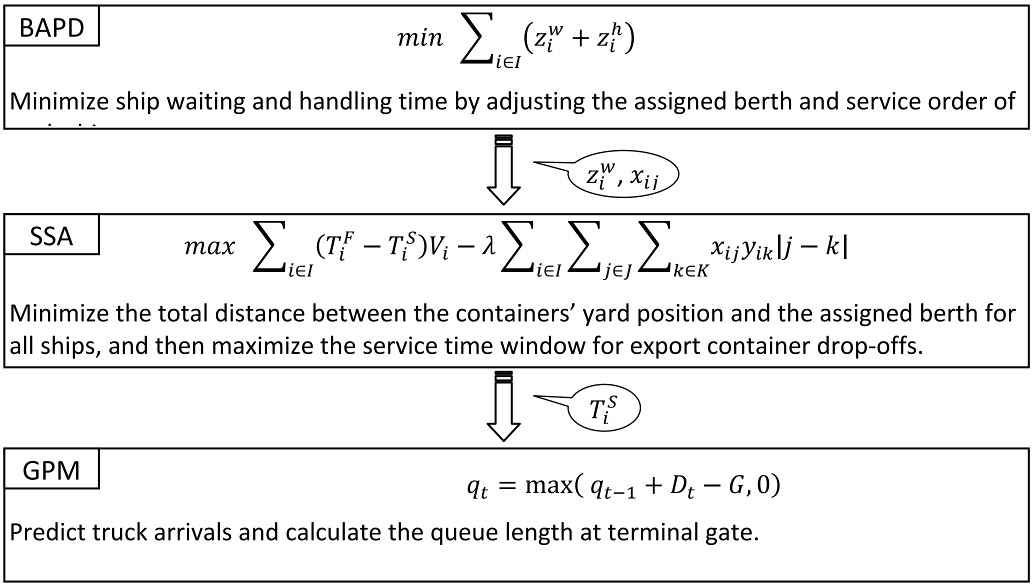
gate time window assignment (GTWA) refers to the problem of assigning vessel-dependent time windows for container drop-offs and pick-ups. BAP is preliminarily solved based on information about estimated gate congestion from GTWA and information about estimated container location from SSA; finally all three preliminary plans have to get confirmed by checking whether yard capacity can afford or not.

In this tactical planning case, where the main concern is to utilize the existing resources in order to improve customer service as high as possible,.
In the rest of this paper, we develop the integrated model by establishing two-way feedbacks between the independent planning parts, then introduce the weekly wrap round effect of the planning into the model.

In a typical maritime container terminal, a set of vessels are scheduled to arrive on a weekly basis. This known seaside demand is used to make the tactical terminal operation planning, basically including a berth template, an associated yard template and an associated gate operation template. These templates are used in turn as the input of the actual

operations, when the vessels and the trucks arrive at the terminal. High quality tactical planning can improve the utilization of the whole terminal system.
The problem of truck congestion at terminal gates has become serious in many major seaports. Currently, we can see two solutions for this congestion problem: gate appointment system in North America and Europe and shared vessel-dependent time window method in China. Nevertheless, the solutions are not integrated into the whole system yet, for example, the vessel-dependent time window method connects only with yard storage planning, not berth allocation planning. Since the operations inside a terminal interact with each other, these isolated solutions seem incapable to fully solve the congestion problem.

and the other ships have to each other. every ship is scheduled either as the first ship or closely following another ship at the same berth.
by adjusting the assigned berth and service order of each ship.



A CHASSIS EXCHANGE TERMINAL TO REDUCE TRUCK CONGESTION AT CONTAINER TERMINALS

Rommert Dekker[1], Sander van der Heide[2], and Eelco van Asperen[3]

[1]Erasmus School of Economics
Erasmus University Rotterdam
Burg. Oudlaan 50
Rotterdam, The Netherlands
e-mail: rdekker@ese.eur.nl

[2]Erasmus School of Economics
Erasmus University Rotterdam
Burg. Oudlaan 50
Rotterdam, The Netherlands
e-mail: sander154@gmail.com

[3]Center for Maritime Economics & Logistics
Erasmus University Rotterdam
Burg. Oudlaan 50
Rotterdam, The Netherlands
e-mail: vanasperen@ese.eur.nl

Abstract: Truck congestion for container terminals seems to be a worldwide unwanted phenomenon. In this paper we present and analyze a chassis exchange terminal concept to reduce the congestion. The terminal works as a kind of extended gate of a group of normal container terminals. During the night containers are collected from the normal container terminals using chassis. In daytime these containers on chassis are then collected and exchanged with export containers on chassis. By exchanging the chassis we avoid extra handling of containers. As connecting and disconnecting to a chassis is done in a short time, the chassis exchange terminal increases handling capacity substantially during peak hours. In this presentation we analyze such a concept for the Maasvlakte container terminals in Rotterdam. We investigate both the effect on waiting time, as well as the environmental effects.

1. INTRODUCTION

World-wide container transport has shown large growth over the past decades and due to ongoing containerization of cargo flows and further globalization this growth is expected to continue in the years to come. To accommodate the volume demand container-ship sizes have shown an increase and will continue to increase. As ocean waterways offer endless capacity it has been recognized that the decisive factor for accommodating and enabling the volume growth will be land-based infrastructural capacity. The availability of sufficient access to deep-sea terminals and hinterland infrastructure (road, rail, waterways) will be the key driver to the success of future container operations.

Although European legislation is strongly advocating to increase the market shares for rail and barge (driven by environmental and congestion arguments) the industry and many legislators at the same time recognize that enforcing a realization of the desired modal-shift will take years.

A lack of available and efficient capacity of these modalities, a strong merchant driven European market (which reduces shipping-lines influence on even realizing this imposed modal-shift), not all ports having inland waterways and the ongoing development of environmental friendly truck technology favors the more practical approach of the co-modality thinking (= optimization within existing modalities) as the best way for achieving an optimal and sustainable logistic system.

Therefore we believe that road-haulage will remain an important mode that deserves and desires further attention and it is for this reason that we like to present an alternative solution that will have a positive impact on road-congestion, sustainability, service integrity and the operational cost of container inland truck operations.

A Chassis Exchange Terminal (CET ©) developed by mr. D. Broek (former executive director of APL Europe) is an off-dock terminal where trucks will exchange chassis instead of unloading and loading containers at the deepsea terminal's premises.

The outline of this paper is as follows. First we describe the typical working of a container terminal, next we outline the CET © solution and thereafter we quantify its effects both from a cost, environmental and efficiency perspective.

2. THE TYPICAL WORKING OF (A IMPORT) CONTAINER TERMINAL

2.1. Description

The main business of a container terminal is the transfer of containers between different transport modes. This can be the transshipment of containers from vessel to vessel, or the link between the hinterland and intercontinental shipment. Hinterland transport of containers can be accommodated by via barge, train and truck. In this article the emphasis will be on the bottlenecks at the truck side of the process. Overviews of the container transport have been given by Vis and de Koster (2003), Notteboom (2004) and Meersmans and Dekker (2001). We will describe the process at a semi-automated container terminal (like the ECT Delta terminal in Rotterdam) from ship to truck in four steps, which is visualized in figure 1.

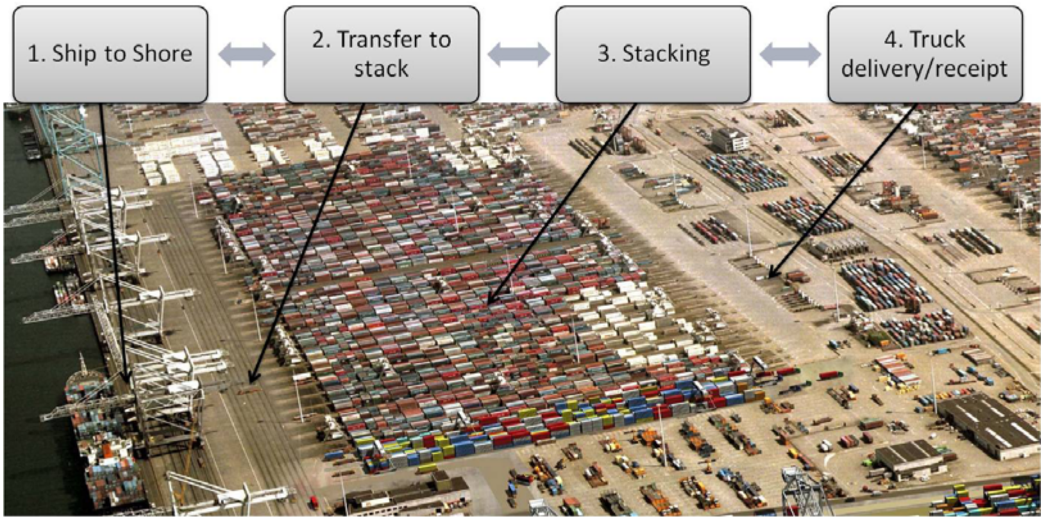


Figure 1. Container Process at the ECT Delta Terminal (Source: ect.nl, 20-06-2010)

The first step shows the transfer of containers from ship to berth using container cranes. During the second step containers are moved from the berth to the stack using AGVs (Automated Guided Vehicles). The third step is the storage of the containers for a certain amount of time. Stacking and unstacking is done via stacking cranes that roll over several rows of containers. The final step is the delivery / receipt of the containers by trucks or intermodal transport. Straddle Carriers transport the containers between trucks and the stack. This process, from step one to four, describes the import of a container. The process will be reversed in case of an export container. Not included in this figure is the gate where trucks enter the terminal. The gate and the processing capacity at the terminals are bottlenecks in the whole truck transport process.

The gate process includes the following: Pre-announcement, check of documents, clearance, trucker gets job, arrives at gate, check of container + trucker, trucker goes to exchange point, unloads container, loads container, trucker goes to gate, customs check, final check, leaves for destination.

2.2. Problems

In Europe 76.6% of the inland transportation of goods is done by trucks (European Union Road Federation (ERF), 2009). This large amount of trucks causes for congestion at terminal gates at peak moments. Container transport is an ever growing sector. This increase will lead to even more congestion at the terminal gates. A study in the Los Angeles and Long Beach ports showed that 40% of the trucks waits for more than 2 hours at the terminal gates, leading to 3.7 million hours of waiting per year for the LA/LB area alone (Barber & Grobar, 2001). Waiting often happens in a queue with the engines of trucks idling. This causes for significant amount of emissions. If no action will be taken it is likely that the waiting times and the emissions associated with this will increase significantly in the future.

Other problems that parties involved in the transport of containers via truck face are:

- Road-Congestion: mainly during rush hours it is likely that truckers are faced with road congestion. Mainly the roads in the harbor area face these problems with both freight and commuter traffic on the road at the same time. This congestion increases the trip time of truckers.
- Dependence on terminal capacity: the turnaround time at seaport terminals is based upon the capacity that can be handled within a certain time and the priority the different jobs on the terminal get. In a merchant haulage situation the transport towards the hinterland via trucks is not the problem of trucks and it will not cost them if trucks have to wait longer. It thus is not in the interest of the seaport terminals to increase their capacity to handle peaks more efficiently.
- Equipment failure: the failure of terminal equipment and especially the IT-systems can cause for unforeseen delays. This is a problem that will only occur sporadic.

3. THE CET

The Chassis Exchange Terminal (CET) is an off-dock container terminal where truckers exchange chassis. A bottleneck trucking companies face is combing an import and export container into one trip within a reasonable turnaround time. The turnaround time increases if a trucker has to visit different sites on a terminal or two different terminals. Instead of visiting a seaport container terminal during peak hours, a trucker exchanges a chassis at the CET. The CET will use off-peak hours to transport the containers from and to the seaport terminals. The CET is based upon two principals, a stacking method called 'stack on wheels' and a shared chassis pool. Stack on wheels is a stacking method where containers are placed directly on a chassis at the terminal. This stacking method requires a lot of space but has as advantage that the handling is reduced to a minimum (Henesey, 2006). Due to the land requirements it is a less used stacking method and only a small number of seaport container terminals in the United States use it. Stack on wheels requires a terminal to have chassis ready for the placement of containers, this is where the shared chassis pool comes in place. The CET accommodates a large chassis pool that visitors of the terminal can use. The CET is intended to be used by a limited number of truck companies which will also be shareholder, in order to avoid cost allocation issues and to increase security of the whole process. One of its shareholders will be a chassis leasing company, which will provide the chassis and which will use the CET for inspection and repairs.

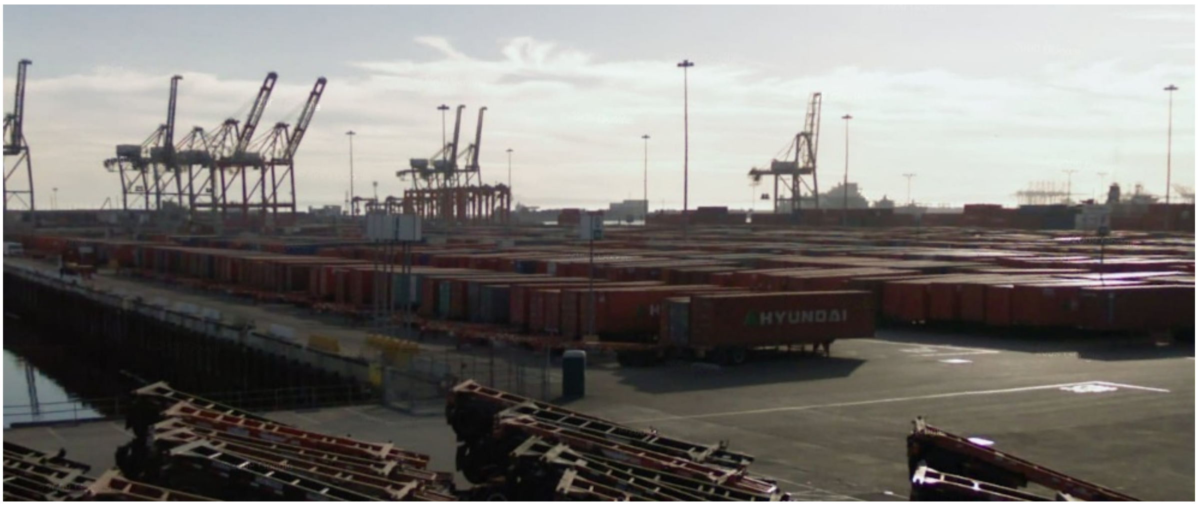


Figure 2. Stack on Wheels at Long Beach, California (Source: maps.google.com (Street view), 27-05-2010)

3.1. The process

The CET has two processes, viz.:

- Chassis exchange: truckers deliver a chassis with an export container and collect a chassis with an import container.
- Shuttle service between CET and seaport terminals: off-peak shuttle service to transfer export/empty containers to the seaport terminals and pick-up import containers for transport towards the hinterland.

The exchange of chassis is a process that will take place during the daytime. Likely the opening hours of a CET will be between 5am and 7pm.

Figure 3 shows the proposed layout of the CET.

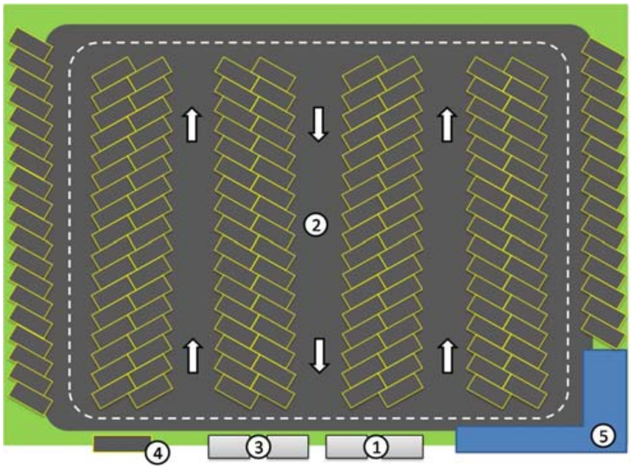


Figure 3. Layout CET (van der Heide, 2010)

The process will be discussed step by step following the numbers on Figure 3.

1. Gate-in: Trucks arriving at the CET will be scanned by the automated gate system. By scanning the truck the system can couple it to an announced visit and guide the truckers to the right location on the terminal. As the truck passes the gate it will receive a ticket with instructions. These instructions include the location to place its current chassis and the location of the new chassis. Truckers that are not recognized by the system are redirected to a parking place (4). This ensures that there are no delays at the gates.
2. Switching chassis: Trucks first proceed to the designated space to uncouple their current chassis. After this is done the truck proceed to the designated parking space where the chassis with the import container is located. Here the new chassis is coupled to the truck and after that the truck can proceed to the gate-out.
3. Gate-out: Trucks arriving at the gate-out will be scanned by the automated gate system. The system checks if the right chassis has been coupled to the truck, if this is the case the truck can proceed toward it hinterland location. If a truck coupled the wrong chassis it is redirected to a parking place (4) to solve the problem.

The entire process is expected to take approximately 15 minutes (Broek, 2009). Due to the fact that the terminal does not rely on any terminal equipment it lacks any bottlenecks involved with this and therefore ensures a fast turnaround time that only depends on the speed of the trucker himself.

The transfer of the containers from and to the CET is done at off-peak hours. Normally these are at night from 21h till 5h. During these hours a pool of trucks will shuttle between the CET and the different seaport terminals. Since the CET has a large number of container to drop-off and pick-up it can easily combine trips to ensure the fastest turnaround times. Unlike trucking companies the CET does not face problems combining trips to the same destination. The number of roundtrips visiting different locations is limited to a small number of dissimilarities in import and export containers or different destinations. By combining trips and using off-peak hours the turnaround time at the seaport terminals will be lower.

3.2. Advantages

The CET is beneficial to several parties within the container transport. The main advantage of the CET is peak shaving. By taking a substantial part of the daily truck visits from the seaport terminal to the CET the peak moments will be quieter. This shift requires that seaport terminals are open at night. The change is beneficial to both visitors of the CET and the seaport terminals. Visitors of the seaport terminal have to share the available capacity with less truckers and thus will experience shorter waiting times. Visitors of the CET have a shorter turnaround time due to the method used that does not rely on terminal personnel or equipment. Accompanied with shorter turnaround times is less waiting and thus idle engine time. This will lead to less emissions.

Besides this advantage the following advantages will occur for parties involved in the CET:

- Increase in daily amount of trips: due to the shorter truck turnaround times at the CET truckers can be back on the road sooner, with this they can possibly avoid the rush hour and they will shorten their total trip time. This will allow them to do more trips on a day compared to truckers visiting the seaport terminals.
- Less empty trips: currently it is common that truckers drive to or from the seaport terminals empty. This is primarily done since it is hard to combine an export and import container without increasing the trip times significantly. The CET offers trucking companies one destination to deliver and pick-up containers. This will increase the utilization of the capacity a truck will have on a round trip. Using a smaller amount of trips for the same capacity will have a positive effect on the environment as well.
- Increase of work load in off-peak hours for seaport terminals: the CET will take a portion of the daily trips away from the peak hours and move this to the off-peak hours. The shuttle transfer between the CET and the seaport terminals will occur during the for seaport terminals usually quiet night. Seaport terminals have tried several methods to make this shift before. Los Angeles for example tried Truck Appointment Systems at first, but since this was not working they introduced a payment system for trucks visiting during the peak hours. This made off-peak hours more interesting (Giuliano, Hayden, Dell'aquila, & O'Brien, 2008).

4. EFFECT OF CET ON GATE CONGESTION

4.1. The problem

Arrivals of trucks at a container terminal vary in time over the days. A typical day pattern of truck arrivals at the ECT Delta is as follows (Voogd, Dekker, & Meersmans, 1999).

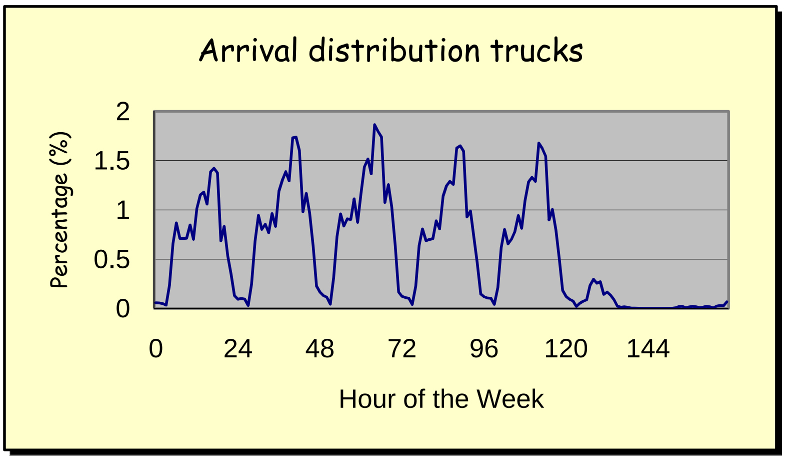


Figure 4. Arrival Distribution of Trucks at The ECT Delta Terminal

This pattern reflects data from before 1999. Nowadays the morning peak is slightly higher. The ECT Delta terminal is closed for trucks on Saturday afternoon and Sunday (ships are still served). During the weeknights it is open, in order to reduce day congestion at the terminal and the connecting A15 highway. Yet few truckers arrive during the night and a project called "Nightdistribution" appeared to be not that successful, despite incentives given. The reason is that truckers mostly cannot deliver their containers during the night. In fact most distribution centers open after 7h and do not accept containers before 8h. Truckers want to work during day-time. They may be willing to start early, even at 5h, but they can only work 12h a day (check** including waiting time). The afternoon peak can be explained from truckers who want to pick-up a container for the next day. They then park their loaded truck at their company and continue the following day. There may be some international truckers arriving at night, but their number is small. This phenomenon is not uncommon in the world, although many terminals are not open during the night. It also depends on the area the terminal serves and the modal split. Export and transshipment terminals suffer less from these problems.

The terminal applies several working shifts to accommodate demand. The shifts ECT operates are (according to their 2009 labor agreement) an early morning shift (from 7.15 – 15.30), a midday shift (from 11.30 – 19.45), an evening shift (15.15 – 23.30) and finally a night shift (23.15 – 7.30). Shifts lasts for 8h15m and have a lunch break of half an hour. Quite often the breaks during shifts increase the congestion. Accordingly it is more difficult for a container terminal to accommodate time-varying demand than e.g. a call center which can work with part-timers. Therefore ECT will not try to accommodate the highest demand rate, but somewhere below that value, also because a container terminal does not have to pay penalties for long delays in handling. A final problem is that demand over the weeks may vary substantially. E.g. the number of trucks may change with 20% over the weeks. Since the economic crises the demand has gone down somewhat as in 2008 they handled even 16,000 trucks a week (Source: ect.nl, 20-6-2010).

It will be clear that because the time varying demand and the difficulties to accommodate capacity to that, as well as the problem of having breaks, there can be substantial queues of trucks. This is a common phenomenon, all over the world. Usually truckers have to accept it, but it will be clear that overall supply chain costs can be reduced if there is much less waiting.

4.2. The effect of the CET

It is not that easy to evaluate the effect of the ECT, as there can be several bottlenecks in the handling. As the CET plans to retrieve containers from the terminal between 19h and 5h, which corresponds with the trough in the truck arrival pattern, it will be clear that the CET allows for a substantial reduction of the peaks. It is however beforehand not clear how the daily pattern will be once the CET has been introduced. The CET is meant for a number of companies in the area of Rotterdam, with the idea that they will be able to do several roundtrips per day. Accordingly it seems realistic that the arrival rate will be lower over the whole day between 5h and 21h.

What will be the effect of a lower arrival rate? It is well-known from queuing theory that even a small reduction of workload can reduce the waiting lines considerably. Presently, some 2000 to 2400 trucks arrive per day at ECT Delta Terminal. It is the idea that the CET will handle some 480 trucks per day which would otherwise go to this (** check: no trucks going to other terminals?) terminal. The effect of such a 20% decrease depends on the average load factor of the Delta terminal, but that varies over the day.

Calculations can be done in two ways, viz. analytical or by simulation. A challenge for the analytical calculations is that the arrival rate varies over the day and surpasses the capacity. One could evaluate the time dependent process or use an approximation. Several approximations are possible in case of the M(t)/M/c queue, i.e. with an exponentially distributed processing time and c servers. Another issue is that the straddle carriers are not the only bottleneck, there is also a gate process and the retrieval from the automated stack area. Processing times of the latter do depend on whether a ship is handled at the same time. Finally processing times also depend on the tasks the truck may have: dropping off and/or picking up a container, and whether it carries or picks-up one 40 or 45 ft container, or 2 20 ft, or a combination.

Below we show the results of simulations on the effect of the CET for both the queue length and the turnaround time at the ECT terminal assuming only the straddle carriers as the main bottleneck.

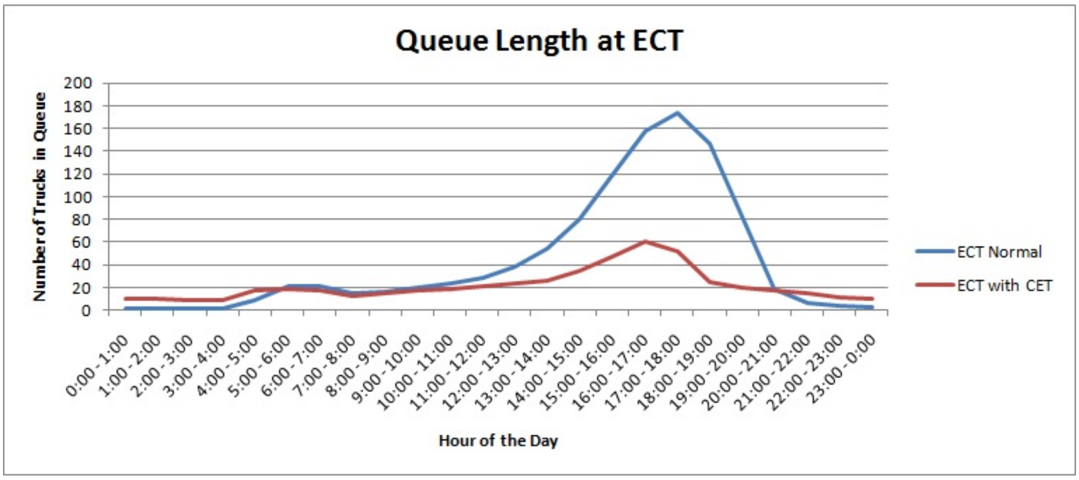


Figure 5. Queue Length at ECT

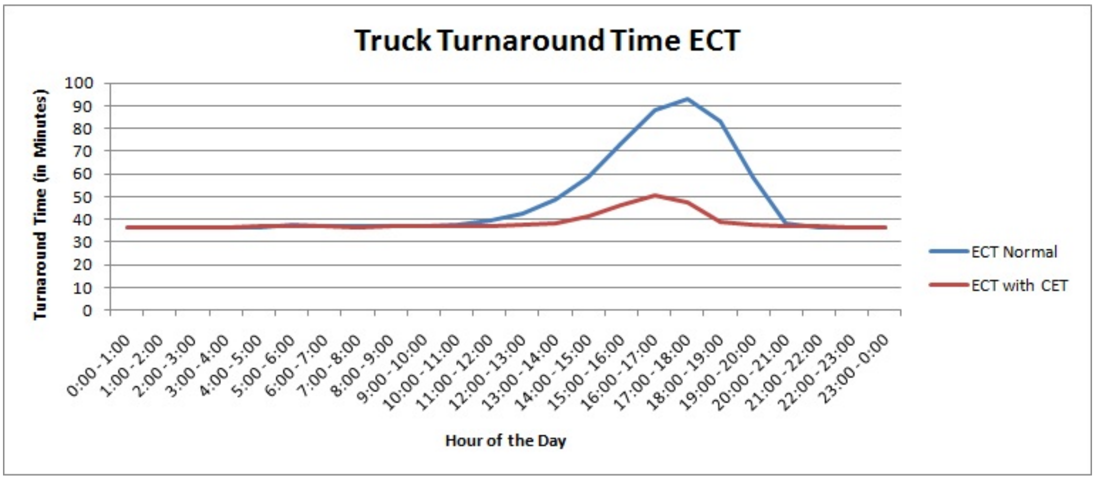


Figure 6. Truck Turnaround Time

The turnaround times will be much less when using the CET. There is one paradox however. If the CET reduces the waiting times substantially, then it will no longer be cost-effective as then all users will avoid the CET as they have to pay for it and not for the standard terminal treatment.

5. EFFECT OF CET ON LOAD FACTOR TO TERMINALS

At the moment there are three terminals at the Maasvlakte: ECT Delta Terminal, the APM Terminal and the EUROMAX terminal. The ECT Delta terminal in fact consists of three terminals, DDN, DDE and DDW, with terminals working for one or more clients. The expansion to Maasvlakte 2 foresees even more terminals: Rotterdam World Gateway, an APM terminal and possible even a sixth terminal.

In order to reduce road congestion and to reduce greenhouse gas emissions, one wants to have high truck load factors. This means that one would like to see trucks making roundtrips loaded on both ways. In general trucks can carry one 40 or 45 foot container or two 20 foot containers. There are some exceptions, the LZV truck that can carry 3 20 foot containers, but its share is at the moment below 0.1%, so we will leave it out of consideration.

Several agencies are monitoring truck load factors, but before going into detail, we have to make some remarks. Western-Europe has a substantial trade imbalance with Asia, this implies that many containers return empty to Asia. For truckers it does not matter whether they transport empty or full containers, yet a truck with an empty container is officially registered as an empty truck, as in law terms the truck with a container is one legal entity and that entity is empty. The next thing is that in the registrations there is no difference between a 20 ft or a 2 x 20 ft loaded truck.

The official truck load factor for trucks moving to the terminals is 60%. The official TEU load of trucks carrying containers back and forth is 2.3 TEU. This implies that from the 4 TEU only some 60% is used. The next thing to consider is the TEU Factor which indicates the ratio of 20-feet and 40-feet containers while talking about a number of TEU. The current TEU factor is 1.7, this means that 70% of the containers is 40-feet and 30% is 20-feet. (thenortherngateway.co.uk, 26-04-2010)

The reason for a low truck load factor is that there is no suitable return trip at the moment the trip is executed, so a triangulation cannot be done (this is a trip followed by a dead trip and next a return trip). So when a container is brought to the Maasvlakte, there is no suitable pick-up container. This may be due to the fact that such a job is only later available, or that it is at another terminal. It will be clear that an exact analysis is very difficult.

6. COST ANALYSIS CET

The use of the CET will cost extra opposed to the seaport terminal. Till now we mainly discussed the upsides of the CET, but do the upsides way up to the extra costs involved. In this chapter we will discuss the costs that are involved with developing and running the CET.

The development of the CET requires a substantial piece of land, the storing of containers on chassis requires a lot more space than other more common stacking method. The location of the CET is close to the seaport terminal, because then a lot of transfers can be made during one night with a limited number of truckers. The land in the seaport area is rather expensive, but locating it much further away would increase the transportation cost substantially. Contrary to traditional terminals the CET does not require special land preparations or terminal equipment. This makes the development cheaper opposed to inland terminal solutions with quaysides or train depots. Required investments are:

- Preparing terrain: pavement of the terrain, assigning parking lots for chassis, fencing and lightning;
- Automated gates system at the entrance and exit;
- Office building for personnel;
- IT hardware and software for terminal operations.

The investment costs needed for the CET are marginal compared to other inland terminal solutions such as train or barge terminals. These inland terminals require ground preparations to support the weight of multilevel stacking and also the construction of a quayside or a train terminal (IBI Group & MacDonald, 2006). These investments are quite expensive and can cost several million. These costs are not necessary for the CET.

For the daily operations the following annual cost will be made:

- Personnel: gates, IT, administration and management;
- Land: rental fee of terminal terrain;
- Terminal maintenance: maintenance, electricity and utilities.
- IT Costs: yearly software subscriptions and depreciation of hardware;
- Chassis pool: rent and maintenance;
- Trucks and truckers for nightly shuttle service.

The visitors of the CET will pay for this in return for the beneficial service offered to them. It is possibly that a cooperation with these parties is done. For example trucking companies could arrange the nightly shuttle service.

7. ENVIRONMENTAL IMPACT OF CET

The CET has several positive effects on the environment. This is done by reducing the terminal load during peak hours and also by increasing the average load factor during round trips.

By shifting trucks from the seaport terminals toward the CET the peaks that seaport terminals face are softened. This lowers the waiting times of truckers and thus decreases the amount of hours that truck engines run idle. The truckers that are going to the CET instead of the seaport terminals are not depending on the terminal equipment and thus are back on the road faster cutting time of their trip and thus exhaust less emission per trip.

The CET gives trucking companies one place to send trucker to for both delivering and collecting containers, where these containers have to go to different seaport terminals. In a normal situation it would be faster and cheaper to split these trips while the CET allows trucking companies to combine these containers to one trip. This increases the use of available capacity and also decreases the amount of trips needed and thus the emissions.

Following is a simulation of the difference that the higher utilization of capacity has on the number of trips, the emissions and also the costs. The simulation uses EURO V trucks, currently the most clean truck type available in Europe, and an average trip distance of 75 km.

Table 1. Emission comparison normal vs CET for 100,000 TEU
Source: NEA Vergelijkingskader Modaliteiten v1.4b (NEA, 2004)

	Normal	CET	Difference
Utilization of capacity	61%	82%	
Loaded trips	85%	95%	
Number of Trips	96,432	64,185	-33.4%
CO2 Emissions (kg)	6,577,279	4,377,817	-33.4%
NOX Emissions (kg)	17,587	11,706	-33.4%
PM10 Emissions (kg)	176	117	-33.5%
Transport Costs	€ 10,546,770	€ 7,387,035	-30.0%
Total Costs	€ 12,539,870	€ 8,713,636	-30.5%

Loaded trips is the fraction of trips on which a container is transported, while utilization measures the fraction of cargo weight versus weight capacity.

By increasing the number of loaded trips and the utilization of the capacity the number of needed trips is reduced by 1/3. This reduction reflects in the emissions as well and the costs are reduced by 30%. The difference in costs allows trucking companies to spend up to 120 euro on a round trip to the CET. Due to its cheaper cost structure it is expected that the CET will be far cheaper than this and thus there will be a healthy profit margin for trucking companies.

8. CONCLUSIONS

The Chassis Exchange Terminal (CET) is a new concept to reduce truck congestion at seaport container terminals and to improve truck efficiency in trips to these terminals. The main idea is to avoid extra crane handling by putting containers on a chassis and applying a chassis pool. The CET is fed from the container terminals during the nights, when these terminals encounter very low truck traffic. The stacking method requires more land, but it reduces the truck turnaround time substantially. Moreover, as there is no real handling bottleneck, it also removes uncertainty in the retrieval of containers, thus allowing truck companies to plan multiple trips from customers to the CET per day. Moreover, it allows for more loaded round trips, as trips to different locations can easier be combined. The higher trip loads will reduce the number of empty trips and thus substantially reduce greenhouse gas emissions.

ACKNOWLEDGEMENT
The authors would like to acknowledge the input from Mr. Broek, who is the inventor of the CET concept and has been the leading person in developing and realizing the concept.

9. REFERENCES

Barber, D., & Grobar, L. M. (2001). *Implementing a statewide goods movement strategy and performance measurement of goods movement in California* METRANS Transportation Center; Available through the National Technical Information Service.

Broek, E. A. (2009). Business plan 'chassis exchange terminal'.

European Union Road Federation (ERF). (2009). European road statistics 2008. *16th International Road Federation World Road Meeting 2009,* Lisbon - Portugal.

Giuliano, G., Hayden, S., Dell'aquila, P., & O'Brien, T. (2008). Evaluation of the terminal gate appointment system at the los Angeles/Long beach ports. *METRANS Project 04-06.*

Henesey, L. E. (2006). Multi-agent systems for container terminal management. (Doctoral, Blekinge Institute of Technology).

IBI Group, & MacDonald, H. M. (2006). *Inland container terminal analysis*British Columbia Ministry of Transportation.

Meersmans, P. J. M., & Dekker, R. (2001). *Operations research supports container handling* No. EI/2001-22)

NEA. (2004). Comparison framework modalities (Computer program in Dutch).

Notteboom, T. E. (2004). Container shipping and ports: An overview. *Review of Network Economics, 3*(2), 86-106.

van der Heide, S. (2010). *Congestion problems at seaport terminals: Comparing innovative solutions, MSc thesis Erasmus University Rotterdam.*

Vis, I. F. A., & de Koster, R. (2003). Transshipment of containers at a container terminal: An overview. *European Journal of Operational Research, 147*(1), 1-16.

Voogd, P., Dekker, R., & Meersmans, P. J. M. (1999). *Famas-NewCon: A generator program for stacking in the reference case* No. EI-9943/A) Erasmus University Rotterdam, Econometric Institute.

GOAL PROGRAMMING APPROACH FOR PORT CAPTIVE HINTERLAND ANALYSIS

Xuejing Yang[1,2], Joyce M.W. Low[3], and Loon Ching Tang[2*]

[1]Automation College, Harbin Engineering University, China 150001
[2]Department of Industrial Engineering, National University of Singapore,
10 Kent Ridge Crescent, Singapore 119260
Email address: g0700998@nus.edu.sg

[3]Centre for Maritime Studies, National University of Singapore,
12 Prince George's park, Singapore 118411
Email address: phan.joan@gmail.com

*Department of Industrial Engineering, National University of Singapore,
10 Kent Ridge Crescent, Singapore 119260
Email address: isetlc@nus.edu.sg
Phone: (+65) 6516 6648
Fax: (+65) 6777 1434

Abstract: This paper proposes a decision tool for shippers in inland cities to optimize port selection problem, and provides a methodology for port operators to identify the captive hinterland of their ports. Specifically, a goal programming model is presented to optimize the port selection problem for cargos originating from inland cities, taking into considerations multiple and conflicting objectives such as minimizing transport cost and time. In the process of optimization, the captive hinterlands of individual ports are identified and delineated. The real-world applicability of the proposed model is illustrated through the various scenarios in 16 Chinese inland cities, in which, results from the model are used to identify the captive hinterlands of three important Chinese ports (Tianjin, Shanghai and Shenzhen).

Keywords: Goal Programming, Port Captive Hinterland, Transport Logistics, Chinese Ports, Hinterland Economy

1. INTRODUCTION

A port's hinterland is the continental area of origin and destination of traffic flows through a port, in other words, it is the interior region served by the port (Klink and Berg 1998). Relationships between ports and the hinterland regions are closely interdependent in their development, function and challenges faced. While hinterland provides a port with its market, ports act as the gateway and nodes linking hinterland to international transport networks. Furthermore, in many instances, the economical, social and cultural aspects of a hinterland are reflected in the characteristics of its serving port(s).

The traditional method to define a port's hinterland is based on the transport rates from the interior region to the port (Mayer 1957). However, the rise of intermodal freight transportation dramatically changed the playing field of the port industry. Under one bill of lading for a door-to-door delivery, the attractiveness of port services depends not only on the basis of its performances, reputations or cost of services, but also on the availability of greater inter-modal coordination which will result in a lower total cost (Yuan *et al.* 2009). Shippers of value-added goods regularly identify reliability and transit time as attributes equal to the importance of affordable freight rates in modal choice decisions (Coyle *et al.* 2003). Meanwhile, a port's hinterland is dynamically changing due to the regional developments. As competitive forces created new linkage possibilities, the traditional definition of port captive hinterlands appears to be oversimplified and outdated.

This paper aims to achieve two objectives. First, it seeks to provide insights on how a shipper select a port for shipment based on multiple and conflicting objectives (such as minimizing general transport cost and time). Second, it answers how a port identifies the captive hinterland of a port with a primary intention of leveraging on this knowledge to design strategies aimed at increasing market share. In order to achieve these objectives, the paper proposes a goal

---

[*] Author for correspondence

programming model that is applied to analyze various scenarios in Chinese inland cities and to identify the captive hinterlands of three important Chinese ports (Tianjin, Shanghai and Shenzhen).

The paper is organized as follows: Section 2 reviews the literature on port captive hinterland and port selection. Section 3 introduces the goal programming model for the identification of the export port for inland cities, based on an optimization of multiple and conflicting objectives such as minimizing intermodal transport cost and time. Section 4 applies the goal programming model to analyze the port selection of various inland cities and the captive hinterlands of Tianjin, Shanghai and Shenzhen ports. Section 5 concludes the study and proposes the future research directions on ports selection and investment.

2. LITERATURE REVIEW

Issues on port selections and port captive hinterland are analyzed from two perspectives, namely, port competitions and the development of ports and their hinterlands. Port competitions have been the subject of much research. Early research on port competitions was qualitative analysis of the factors affecting port competitiveness, and found that service-related components were more important than price factors in influencing the competitiveness of a port. Subsequent studies concentrated on economic models of port selections from the viewpoint of the carriers. Hoyle and Charlier (1995) illustrated the historical processes involved in the development of an integrated and concentrated port hierarchy, as well as, highlighted the problems of inter-port competition based on the East African port system. Nam (1997) proposed a discrete choice model for carriers' freight transportation decision, and examined if aggregation over heterogeneous commodity types can cause a bias in model estimations. Cullinane and Khanna (2000) presented a model to quantify the economics of scale in operating large containerships and analyzed the impact of the ever-increasing containership size on the container operations, logistical systems and ports. Following, Malchow and Kanafani (2001 and 2004) proposed a discrete choice model to analyze the distribution of maritime shipments of four commodities among ten US ports, and found that location is the most significant characteristic of port competitiveness. More recently, Yeo *et al.* (2008) identified and evaluated the competitiveness of container ports in Korea and China, and presented a structure for evaluating the components influencing port competitiveness. Tongzon (2009) contributed a study to evaluate the major factors influencing port choice from the Southeast Asian carriers' perspective. In contrast to Malchow and Kanafani, Tongon found that efficiency is the most important factor, followed by shipping frequency, adequate infrastructure and location. From the network perspective, Tang *et al.* (2008) introduced a network-based choice evaluation (N.I.C.E) model and determined that port efficiency and scale economies are the more important dimensions influencing liner shipping companies' selection of major Asian ports. Nevertheless, Tang *et al.* highlighted the importance of a port being an all-rounder in order to ensure competitiveness.

Amidst the literature above, Malchow and Kanafani (2004), and Yeo *et al.* (2008) are the only works that quantitatively investigated the problem of port selection and port captive hinterlands. To the best of our knowledge, none of the extant literature has focused explicitly on the quantitative analysis on ports selections and the captive hinterlands of Chinese ports, which represents 22.3 percent of world container port throughput in 2008 (UNCTAD 2009).

3. MODEL FORMULATION

In an international transportation market, a common objective for both the shipper and carrier is to transport each shipment from its origination to its destination as efficiently as possible. The level of efficiency is defined by transit time and cost. The ports that are being chosen are those that will help to minimize the total sum of sea, port and inland costs and transit time.

Nonetheless, port charges, ocean freight rates and intermodal transfer time are omitted in the model formulation. Some justifications for their omissions are available. As noted in Tongzon (1995) and Murphy *et al.* (1991), port charges are relatively insignificant since they constitute small proportion of total transport costs and shippers are willing to pay higher port cost in exchange for superior services. Similarly, ocean freight rates are not very significant in port selection as Wilmsmeier *et al.* (2006) found that oceanic transportation costs only increase by 29.5 percent when distance doubles. Although intermodal transfer time is highly dependent on various carriers at each port within the choice set, intermodal transfer time for a particular shipment's is generally known with little uncertainty prior to port selection. We deem that it would be advisable to exclude it from model to avoid further complication. As such, the goal programming model will incorporate only the decision variables pertaining to inland transport time, oceanic transport time and inland transport cost.

The proposed goal programming model takes from a shipper's perspective. The shipper desires to select exporting ports for a full 20 feet container from some specified inland cities to some specified off board ports that best optimizes on the trade-off between total transit time and inland transport cost. The concept of the model is illustrated in Figure 1.

We impose several assumptions to avoid unnecessary structural complications in our analysis. First, shippers have complete information about the market and can freely select ports for importing and exporting. Second, the transportation capacity of each link in the model can be ignored because an overloading on a link will lead to congestions that adversely influence its competitiveness in the form of longer transit time. Third, shippers are not faced with scheduling constraints and time window problem when optimizing their objective functions. Henceforth, we could analyze the effect of two strategies, namely, (1) minimize the distribution costs incurred in land transport; (2) minimize inland transport and oceanic transport time, on port hinterlands. Notations that are used in the model can be classified into index sets, parameters and decision variables as follows:

Index sets

I Set of all inland start nodes

J Set of all inland end nodes

O Set of origin nodes

D Set of off board ports

E Set of export ports

OL Set of inland nodes directly linked from origin

LD Set of inland nodes directly linked to export ports

Parameters

W_1 Weight determined by relative importance of less transit time

W_2 Weight determined by relative importance of lower costs

C_{ij} Inland freight rate from i to j

T_{ij} Inland transit time from i to j, expect value based on major carriers

O_{ij} Oceanic transit time from i to j, expect value based on major carriers

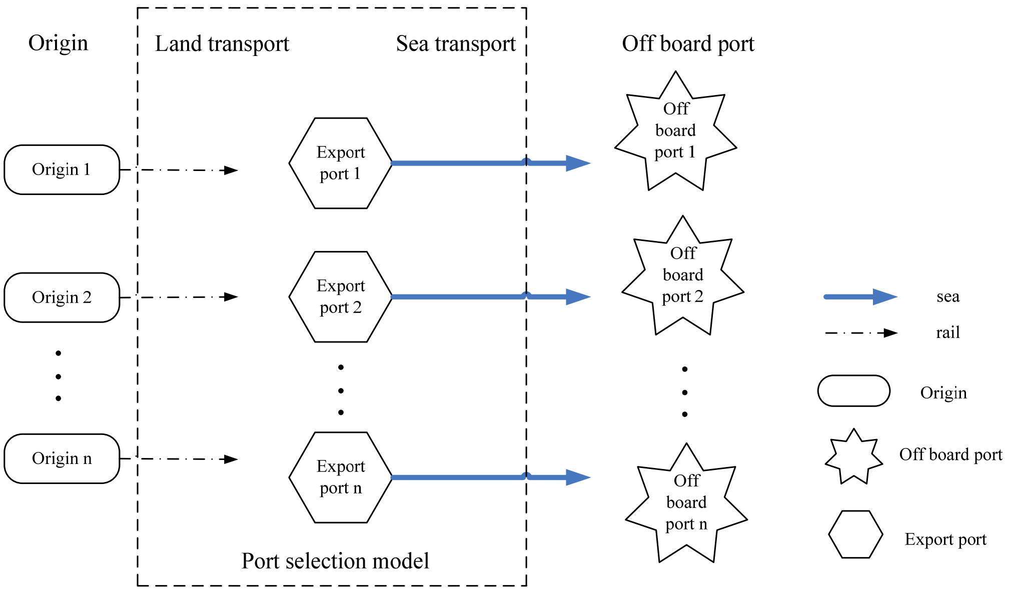


Figure 1. Concept of The Model

Model

$$\text{Min } Z = W_1 d_1 + W_2 d_2 \quad (1)$$

Subject to:

$$\sum_{i\in I}\sum_{j\in J} T_{ij} x_{ij} + \sum_{i\in E}\sum_{j\in D} O_{ij} x_{ij} - d_1 = 0 \quad (2)$$

$$\sum_{i\in I}\sum_{j\in J} C_{ij} x_{ij} - d_2 = 0 \quad (3)$$

$$\sum_{j\in OL} x_{ij} = 1 \qquad i \in O \quad (4)$$

$$\sum_{i\in LD} x_{ij} = 1 \qquad j \in E \quad (5)$$

$$\sum_{j\in D} x_{ij} = 1 \qquad i \in E \quad (6)$$

$$\sum_{i\in E} x_{ij} = 1 \qquad j \in D \quad (7)$$

$$\sum_{i\in I} x_{ih} = \sum_{j\in J} x_{hj} \qquad h \neq i \neq j,\ h \in I - O \quad (8)$$

$$\sum_{i\in LD} x_{im} = \sum_{j\in D} x_{mj} \qquad m \neq i \neq j,\ m \in E \quad (9)$$

$$x_{ij} = 0 \qquad \text{if link } (i, j) \text{ is not an available link} \quad (10)$$

Decision variables

X_{ij} Binary decision variable: 1, if link (i,j) is included in the selected route; 0, otherwise

d_1 Positive deviational variable that represents total transit time

d_2 Positive deviational variable that represents inland transport costs (in US$)

The objective function (1) minimizes the weighted sum of deviations from the lowest possible inland distribution cost and the lowest possible total transport time, with priorities represented by W_1 and W_2. Constraints (2) and (3) are referred to as the goal constraints. Constraints (4)-(10) are the feasibility constraints. Constraint (4) ensures every inland origin node has only 1 outflow. Constraint (5) ensures all feasible links to export port has only 1 inflow from the set of inland nodes. Constraint (6) ensures every export port has only 1 outflow to a set of off board ports. Constraint (7) ensures all feasible links to each off board port has only 1 inflow from the set of export ports. In other words, constraints (6) and (7) define a one-one relationship for every pair of export and off board ports. Constraints (8) and (9) are flow conservation constraints that ensure flow continuity for all the inland transit nodes and export ports. Constraint (10) rules out the infeasible links and defines the structure of the transport network.

4. PORT SELECTION: CHINA CASE STUDY

This section applies the preceding port selection model to analyze how 16 major Chinese inland cities choose their export ports from the three most important port cities (Tianjin, Shanghai and Shenzhen) along the Chinese coastline, from north to south. The 16 inland cities were distributed in northeast, central, west, southwest and southeast parts of China. The inland cities considered include Harbin and Shenyang in the northeast region, Beijing, Taiyuan and Zhengzhou in the central region, Xian, Lanzhou, Chengdu and Chongqing in the west region, Guiyang, Kunming and Nanning in the southwest region, and Nanchang, Changsha, Wuhan and Hefei in the southeast region.

The three important port cities are the centers of the most prosperous regions in China. As China's economy driven engines, Shanghai located in the region of the Yangtze River delta, Shenzhen located in Pearl River delta and Tianjin located in Bohai bay. The container throughput of the three ports in 2008 are 27,980,000TEU (Shanghai), 21,413,888TEU (Shenzhen) and 8,500,000TEU (Tianjin) respectively, among world top 20 container ports in the last five years (UNCTAD 2009). Since China's top two export destinations in the last ten years are United States and Europe, we choose Singapore as the off board port for Chinese cargos on Asia-Europe line, and Los Angeles as the off board port for Chinese cargos on transpacific line.

In order to simplify analysis on how the 16 inland cities choose from Shanghai, Tianjin and Shenzhen as their export ports for cargos to Singapore and Los Angles, we only consider rail container transport for inland transport from the 16 cities to export ports. Since the distances from the 16 cities to Shanghai, Tianjin and Shenzhen are all more than 300 km, rail transport is the major mode for container shippers. Our model was solved using Solver 2007 software on Pentium Dual Core T2060. The following analysis on export port selection will be presented based on two destinations: Singapore and Los Angeles.

4.1 Port Selection for Shipments to Singapore

The best solutions for cargos from the 16 inland cities to Singapore are illustrated in Figure 2. For the short oceanic and land distance, Shenzhen is found to be the best export port for all the cities in west (Xian, Chongqing, Lanzhou and Chengdu), southwest (Kunming, Nanning and Guiyang) and southeast (Changsha, Wuhan and Nanchang) regions of China to Singapore. Shanghai is the best choice for cargos from cities in east region (Hefei). While for cities in the central part of China (Taiyuan and Zhengzhou), Shanghai and Shenzhen are almost as good as each other in terms of transport time and cost. Therefore, shippers would make decision base on other elements such as port charges and intermodal transit time. Although Tianjin is the closest port for central city Beijing and north east city Shenyang, Shenzhen would be a better choice for these two cities as both cities have fast and convenient rail transport to Shenzhen which dramatically reduce the total transport time to Singapore. Despite the fact that the oceanic transport time from Tianjin to Singapore will need four more days than from Shenzhen, Tianjin is the best export port for cargos with low time value originating from northeast city Harbin.

For the analysis of commodities with high time value, all the solutions remain the same as before. The only city that would be better off by choosing Shanghai as export port instead of Tianjin is Harbin, as doing so can reduce three days of oceanic transport time.

4.2 Port Selection for Shipments to Los Angeles

Figure 3 illustrated the best solutions for cargos from the 16 inland cities to Los Angeles. Since the oceanic distances from Shenzhen and Shanghai to Los Angeles are similar, Shanghai is the best option for more inland cities in this condition. For cities in southwest (Guiyang, Kunming and Nanning), Shenzhen is the best option for cargos to US, while Shanghai is still comparable good for these cities as the second best choice. For cities in west (Chongqing, Chengdu, Xian and Lanzhou), Shenzhen port is the first choice for Xian and Chengdu, and Shanghai port represents the next best alternative. In the case of Chongqing and Lanzhou, either Shenzhen port or Shanghai port can be the best solution. For cities in southeast and east region, Shanghai is the best solution for Changsha and Hefei. Both of Shenzhen and Shanghai ports offer the optimal solution for Wuhan and Nanchang. Unlike the previous cases, Tianjin, Shanghai and Shenzhen ports are all as good as one another for cities in central region (Beijing, Zhengzhou and Taiyuan), while Shanghai is slightly better for Taiyuan and Zhengzhou than the other two ports. Meanwhile, for cities in northeast region, Shanghai is the best option for Shenyang, and Tianjin is the best option for Harbin.

For commodities with high time value, all the solutions remain the same as displayed in Figure 3. Harbin continues to be the only inland city in which Shanghai port should be chosen as export port over Tianjin port. Shanghai port offers a shorter oceanic transport time by three days.

4.3 Ports Captive Hinterland Analysis

Based the analysis on port selection of the 16 Chinese inland cities, the captive hinterlands of Shanghai, Shenzhen and Tianjin ports can be delineated. Shenzhen, as the center port city in the Pearl River delta, has effectively captured broad regions in southwest, south, west and the central part of China. The port is especially attractive for commodities with high time value originating from these regions and heading to the European markets. The convenient rail transport to inland cities and the speedy oceanic transport to Singapore strengthened Shenzhen's dominate position as an export port on Asia-Europe line. Shanghai is the hub port for many inland cities in the southeast, east, central and west regions. Especially for cargos to U.S. market, Shanghai and Shenzhen are comparably as good as each other for inland cities in central and west regions. Tianjin port indicates inferior competitiveness than Shenzhen and Shanghai in the preceding analysis. The main reason would be that many carriers' shipping schedule for Asia-Europe line and transpacific line do not have direct oceanic transport from Tianjin to Singapore or Los Angeles, and most of the vessels leaving Tianjin would berth at Shanghai or Shenzhen before shipping to Singapore and Los Angeles, which results in a delay of at least three or more days. It can, hence, also be inferred that the shipping demand of Bohai bay rim would be less than Yangtze River delta and Pearl River delta. However, with the rapid economy development of Bohai bay rim, the captive hinterland of Tianjin is expected to be enhanced and extended in the near future.

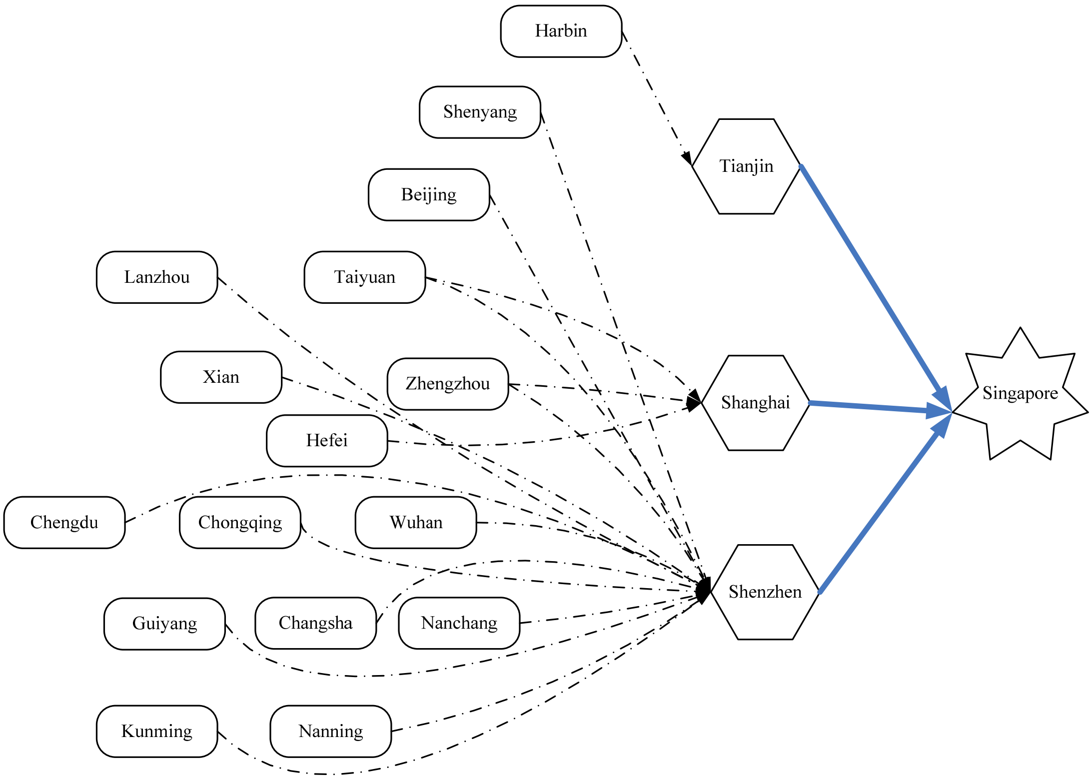


Figure 2. Port Selection for Cargos to Singapore

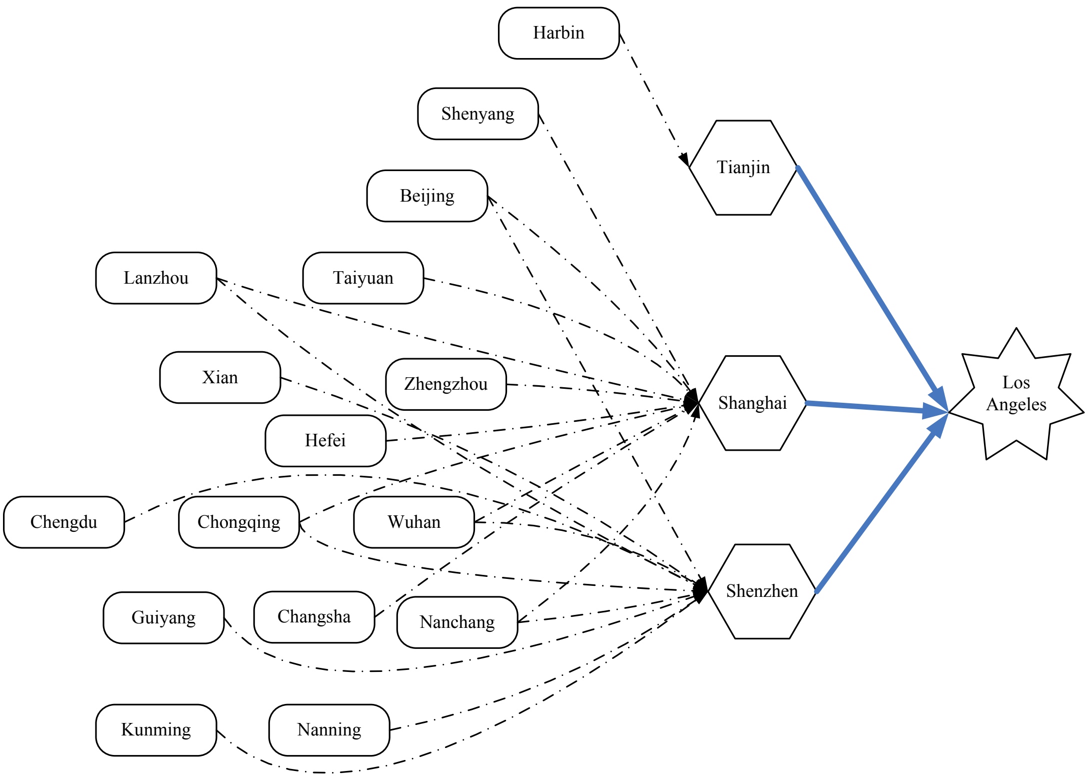


Figure 3. Port Selection for Cargos to Los Angeles

5. CONCLUSION

This paper proposes a goal programming model to optimize port selection problem for inland cities and identify the port captive hinterlands, taking into account of the conflicting objectives in transport cost and time minimization. Based on the comprehensive data sourced from industry, a China case study is conducted to analyze how 16 inland cities choose from three prominent Chinese export ports for cargos with varying time sensitivity heading towards the European and US markets.

Findings from this paper highlight four important aspects. First, inland freight rates would be a less critical factor compared to total transport time in port selection. Second, while the distance from inland cities to the export port would be an insignificant element in itself, ports with convenient inland transport connection and offer less total transport time would attract cargos from far in land cities. Third, a port may enhance or extend its hinterland by promoting in-port service efficiency and speeding up inland and oceanic transport. Future examination could be applied on the consideration of transport volume limitation and its impact on port selection.

6. REFERENCES

Coyle, J., Bardi, E. and Langley, J. (2003) The Management of Business Logistics: A Supply Chain Perspective, 9th ed., South-Western, Mason, OH.

Cullinane K. and Khanna, M. (2000). Economies of scale in large containerships: optimal size and geographical implications, *Journal of Transport Geography* 8: 181-195.

Hoyle, B. and Charlier, J. (1995). Inter-port competition in developing countries: an East African case study, *Journal of Transport Geography* 3(2): 87-103.

Malchow, M. and Kanafani, A. (2001). A disaggregate analysis of factors influencing a port's attractiveness, *Maritime Policy & Management* 28(3): 361-373.

Malchow, M. and Kanafani, A. (2004). A disaggregate analysis of port selection, *Transportation Research Part E* 40: 317-337.

Mayer, H. M. (1957). The port of Chicago and the St. Lawrence seaway, *Department of Geography research papers* 49. University of Chicago.

Murphy, P. R., and Dalenberg, D.R. (1991a). Selecting Links and Nodes in International Transportation: An Intermediary Perspective. *Transportation Journal* 31(2): 33- 40

Nam, K. (1997). A study of the estimation and aggregation of disaggregate models of mode choice for freight transport. *Transportation Research E: Logistics and Transportation Review* 33(3): 223-231.

Tang, L.C, Low, J.M.W and Lam, S. W. (2008). Understanding port choice: a network perspective, *Networks and Spatial Economics*. DOI 10.1007/s11067-008-9081-8.

Tongzon, J. (1995). Determinants of port performances and efficiency, *Transportation Research Part A* 29(3): 245-252.

Tongzon, J. (2009). Port choice and freight forwarders, *Transportation Research Part E* 45(1):186-195.

United Nations Conference on Trade and Development Secretariat (2009). Review of Maritime Transport 2009.

Wilmsmerier, G., Hoffman, J. and Sanchez, R.J. (2006). The impact of port characteristics on international maritime transport cost. *Research of Transport Economy* 16: 117-140.

Yeo, G., Roe, M. and Dinwoodie, J. (2008). Evaluating the competitiveness of container ports in Korea and China, *Transportation Research Part A* 42: 910-921.

OPTIMIZING AND SIMULATING LOADING/DISCHARGING RESOURCE ALLOCATION IN A MULTIMODAL CONTAINER PORT TERMINAL

E. Zehendner, G. L. Rodriguez-Verjan, N. Absi, S. Dauzère-Pérès*, D. Feillet

Department of Manufacturing Sciences and Logistics, Centre Microélectronique de Provence
Ecole des Mines de Saint-Etienne
880 Avenue de Mimet
13120 Gardanne, France
e-mail: zehendner@emse.fr, grodriguez@emse.fr, absi@emse.fr, dauzere-peres@emse.fr, feillet@emse.fr

Abstract: This article is based on a study of a container terminal of the *Grand Port Maritime de Marseille* (GPMM) in France. The goal is to optimize the allocation of the critical loading/discharging resources shared by vessels and inland transport modes (road, rail and waterways). After a short description of the context and a literature review, an integer linear programming model, based on a network flow representation of the problem, is introduced,. This model being deterministic, a discrete event simulation model is proposed to validate the solutions determined by the mathematical programming model in a stochastic setting. Numerical experiments are reported and discussed for both models.

1. INTRODUCTION

Seaports and container terminals are confronted with multi-trade strategic carrier alliances with strong bargaining power. This powerful demand side causes strong competition between container terminals. Traditionally, a port's competitiveness is based on the vessel turnaround time combined with low rates for loading and discharging. However, the competitiveness of a container terminal is becoming more and more a function of its delays and costs for transport to and from the inland. Inland transportation costs make up a large part of total container shipping costs. The European Sea Ports Organization estimates that inland transportation accounts for 40% to 80% of total shipping costs (e.g. ESPO 2004). Inland logistics offer thus important possibilities to reduce overall transportation costs. Moreover, geographical close ports serve similar inland areas. Hence, terminals offering an efficient connection to the inland have an advantage over other terminals, as containers arrive earlier at their final destinations. Customers also are very sensitive to reliable delivery times. Finally, environmental aspects are another reason for the increasing interest on landside logistics. A better organization and coordination of inland transportation should reduce truck idle times at ports and congestion in and outside ports. Therefore carbon dioxide emissions in the port area could decrease.

The *Grand Port Maritime de Marseille* (GPMM) in France is a general cargo port dealing with various types of traffic including crude oil and oil products (oil, gas and chemical products), general cargo (containers and other packaging), dry bulk (minerals and cereals) and liquid bulk (chemicals and food), as well as passenger traffic. GPMM is among the biggest ports in France and in Europe in terms of total throughput. It is especially active in oil cargo and is the third oil port worldwide, but plays only a minor role in the transportation of containerized goods in Europe and in the Mediterranean area. With the objective of increasing its volume of containerized cargo, new container terminals are constructed. GPMM and the container terminals are aware of the importance of a fast and reliable connection to the inland to attract shipping lines. The delays for inland transportation should be decreased and its reliability increased by optimizing the service of trucks, trains and barges at container terminals.

At present, priority is given to vessels. Trucks, trains and barges are unloaded and loaded with the remaining capacity. This may lead to long waiting times for trucks and barges. An optimized resource allocation of internal transport resources (straddle carriers) to vessels and inland transport modes should help to reduce waiting times. This paper aims at introducing an optimization model to determine an efficient resource allocation for a given workload, which minimizes the time trucks, trains and barges spend at the terminal. All data for this analytical allocation problem is assumed to be deterministic and known in advance, although this is clearly not the case at a real terminal. Egbelu (1987) showed that non-simulation approaches underestimate the vehicle fleet size. The validity and robustness of the solutions provided by our optimization model in a stochastic environment is tested via discrete event simulation. The simulation model is a statistically validated representation of a real container terminal at GPMM. The optimization and the simulation models may be used to compare different scenarios at operational, tactical and strategic levels.

2. SHORT LITERATURE REVIEW

The literature related to container terminals is extensive. It includes various areas of studies such as berth and quay crane allocation, storage space allocation and stacking, allocation of yard equipment and yard routing. Steenken *et al.* (2004) and Stahlbock and Voss (2008) provide a detailed literature review of operations research models and their applications to container terminals. Due to the focus on ship turnaround time as a terminal performance measure, most of the literature focuses on the organization of resources and procedures in order to minimize the vessel service time. Few papers are concerned with resource allocation with the objective to minimize the delay of trucks, trains and barges. Some studies focus on minimizing truck turnaround time, but do not address the internal resource allocation problem. They analyze the problem from the haulage contractor's point of view (e.g. Namboothiri, 2006) or the impacts of appointment systems (e.g. Huynh and Walton, 2008).

Vis *et al.* (2005) propose an integer linear program minimizing the number of vehicles required to transport containers between the quay and the yard. Gambardella *et al.* (2001) deal with the problem of allocating quay cranes, transport vehicles and yard cranes to vessels with regard to their workload. The problem is modeled as a network design problem using a mixed integer linear program. Our study is quite similar to theirs. In both cases, the resource allocation problem is modeled as a network problem where arc capacities are limited by the number of allocated resources and containers to be moved are modeled as flows. The feasibility of the obtained solution is verified with a simulation model of a real terminal that is described in Gambardella *et al.* (1998). Their goal is to compute the cheapest resource allocation to serve vessels, whereas we include all transport modes (maritime, road, rail and waterways) in our analysis and aim to minimize the waiting times for landside transport modes.

Bish *et al.* (2005) study the problem of dispatching vehicles to containers for the transportation between quay cranes and storage positions to minimize ship turnaround time. Hartmann (2004) develops a general model to assign jobs to container terminal resources and to temporally arrange the jobs with regard to precedence constraints and sequence-dependent resource setup times. He presents one example where the general model is used to schedule container movements between quay, yard and landside gates on straddle carriers. The objective is to serve quay cranes and inland transport modes at time and to reduce empty travel of straddle carriers. The objectives of his study are very similar to ours. But in his work tasks have to be scheduled on a given set of resources, whereas our work aims to determine the size of this set of resources. Sanchoy and Spasovic (2003) focus on container terminals where, as in our case, straddle carriers are primarily used to move containers. They develop an assignment algorithm that dynamically matches straddle carriers to waiting trucks. The objective is to minimize truck serving times and empty travel of straddle carriers. They also carried out experiments to see the impacts of a reduced number of straddle carriers on the performance.

The previous articles focus on the dispatching of a single resource. Other studies as Froyland *et al.* (2006), Chen *et al.* (2007) and Lau and Zhao (2008) present models for task scheduling explicitly considering the interdependency of different equipment like quay cranes, yard vehicles and yard cranes.

The problem studied in this work is motivated by the terminal operator's wish to improve its inland connection and hence its overall competitiveness. The proposed mixed integer linear problem is based on the situation at a container terminal of GPMM. Its solution may help the terminal operator to decide how many resources should be ordered for the next day. It also indicates how these resources could be allocated to landside and maritime transport modes in order to minimize delays. A simulation model representing the situation at the terminal at GPMM has been developed and statistically validated. It is used to verify the robustness of the solutions of the optimization model in a stochastic environment. Both models may be used to compare different scenarios at operational, tactical and strategic levels. For example, the impacts of different service strategies or of an increasing volume of containers arriving on barges may be analyzed. From a scientific point of view and to the best of our knowledge, none of the existing work in the literature has studied the resource allocation problem with the objective to minimize the service of landside transport means like trucks, trains and barges. Several studies consider the resource allocation problem with the aim to minimize vessel turnaround time, but landside transport modes' waiting times are neglected. It is not possible to transfer seaside resource allocation models to the landside without adjustments, because of differences between vessels and landside transport modes in terms of cargo volume and frequency and certainty of arrival.

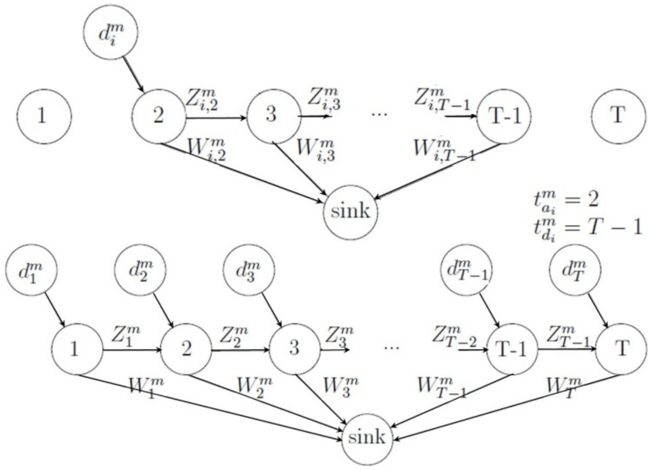
3. AN INTEGER LINEAR PROGRAMMING MODEL

Different transport means (such as vessels, trucks, trains and barges) with different characteristics deliver and pick up containers at container terminals. These transport modes differ with regard to cargo volume, operating costs, associated traffic route and travelled distances, knowledge and reliability of arrival dates, due dates and required handling equipment. Each transport mode is served with a specific strategy which takes into account the transport mode's specific requirements. The penalized waiting times may also differ between different transport modes. Either the delay of a vehicle (truck, train, barge or vessel) or the delay of a task (container movement request) may be penalized. If a vehicle requires only one container movement, the two cases are identical. Container terminals may have different organizations and priorities, thus service strategies and delay costs for the same transport mode may differ from terminal to terminal.

The different transport modes can be served quite independently. They are only related by the fact that they have to share the available resources. Therefore, independent models can be formulated for each transport mode considering its specific characteristics. To represent a real container terminal, we combine the transport mode specific models. These models are only related by one constraint which is the number of available resources at the terminal. The modular structure of the model allows representing a variety of different terminals. To take priorities between different transport modes into account, some transport mode specific delay costs may be added. As seen above, the delay of a task or a vehicle may be penalized. Therefore, the transport mode specific costs should also take this difference into account.

The integer linear programming model calculates how many resources should be allocated to each vehicle of the different transport modes per period. Each vehicle is represented via its arrival time and the number of tasks that have to be executed for it. The vehicle may leave the terminal when all its tasks have been executed. The aim is to minimize the time the vehicle spends at the terminal. The problem is modeled as a network flow problem where the different periods are related by the number of tasks delayed from one period to another. The number of required and executed tasks is compared at each period to decide whether the vehicle may leave the terminal or not. Under some circumstances, it may be possible to aggregate the tasks of several vehicles of the same transport mode to one demand arriving over time. This demand is calculated for each period by summing up the tasks of all vehicles arriving during this period. Such an aggregation is possible if resources may be shared among vehicles and if all vehicles have the same delay costs. In this case the optimal solution for the aggregated problem can be transformed in an optimal solution of the original problem with a greedy algorithm like smallest vehicle first or earliest due date first. First experiments conducted on a few instances solved using a commercial solver show that aggregating vehicles reduces the problem size and accelerates the solution process.

The different models for each transport mode are based on the network flow model. Therefore, the basic structure of all models is identical, but additional parameters, constraints and variables have to be added to each model in order to represent the characteristics of each transport mode. Each network flow model consists of a set of vertices and arcs. The vertex set consists of one or several sources, one sink and several period nodes representing the discrete time periods of one working day. Each source is connected to exactly one period node. The flows on these arcs represent the arrival of tasks at a given period. All period nodes are connected to the sink. The flows on these arcs represent the executed tasks per period. The capacities of these arcs are limited by the number of allocated resources at each period. Period nodes are also connected to their neighbors. The flows on these arcs represent the unexecuted tasks per period which are transferred to be served during the next period. This formulation has similarities with Gambardella *et al.* (2001) where the tasks are also represented as flows in a network and the arc capacity depends also on the number of allocated resources.

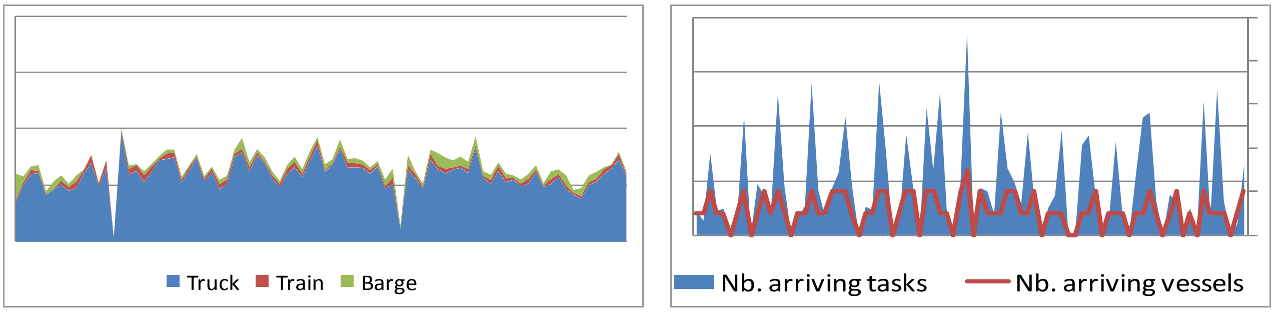


(a) For a single vehicle (b) For an aggregated demand

Figure 1. Scheme of The Network Flow Model

The flow network representing one vehicle has exactly one source. In this case, tasks arrive only once with the vehicles arrival. For each vehicle, a flow network has to be solved. As seen before, vehicles should be aggregated whenever possible. In this case, the flow network has several sources. Each source represents the sum of all tasks arriving on different vehicles per period. Figure 1 shows the two flow networks for a single vehicle and for an aggregated demand. For conciseness, the mathematical model will not be presented in this abstract. This model is presented in Zehendner (2010).

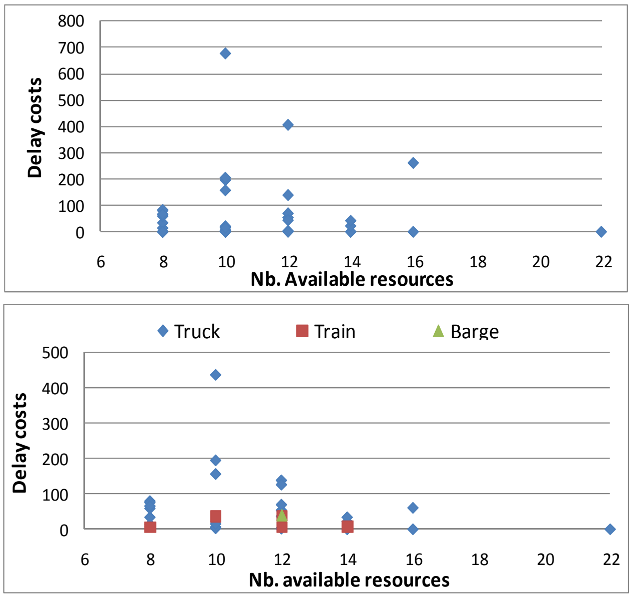
The model has been adapted to the case of one of the container terminals at GPMM (*Grand Port Maritime de Marseille*) in France, taking into account specific constraints of each transport mode. Some of these constraints are described in the sequel. Barges are served at the same quays as vessels. They should be unloaded and loaded as fast as possible. Moreover, barges have to be served prior to a given due date so as to free the quay equipment for vessels. Resources are allocated to exactly one barge and are not shared among barges. Vessels have to be served during their imposed time windows, but there is no incentive to finish earlier. Resources are allocated to exactly one vessel and are not shared among vessels. Trains are served at the rail station. The railcars stay during a fixed time window at the terminal and are picked up by an engine according to a fixed schedule every day. Penalty costs have to be paid for every container that remains at the terminal when the train leaves. Containers are transported between the yard and the buffer in front of the railcars by straddle carriers. The loading and unloading to and from railcars is then done by reach stackers. Straddle carriers are not allocated to a single train, but to trains in general as they only do the transport from the common buffer to the yard and vice versa. Trucks are assigned to parking slots where they are loaded and unloaded directly by straddle carriers. Trucks should be served as fast as possible. Given the small number of tasks per vehicle, it is not efficient to dedicate one resource to one vehicle. Resources are thus shared among several trucks.



(a) Landside Transport Modes (b) For an Aggregated Demand

Figure 2. Number of Tasks To Be Executed Per Day for Different Transport Modes

We obtained actual data from one terminal at GPMM, indicating the number and times of handled import and export containers per transport mode over a period of six month. For confidentiality reasons, some figures, such as the number of containers that are discharged from and loaded to the landside transport modes, cannot be given. On average, export containers make up 48% of total landside container traffic and import containers 52%. The landside transportation for import and export containers is mostly done by trucks (89%). Only a relatively small percentage is transported via barge (6%) and trains (5%). On average 1.3 vessels arrive per day with up to 1100 container movement requests per vessel. Figure 2 shows the number of handled import and export containers for landside transport modes and the number of vessels arriving per day with their required container movements (the scales on the Y-axis have been hidden). 22 straddle carriers are operational at the terminal at GPMM. The number of straddle carriers in use is variable, as straddle carriers have to be driven by operators.



(a) Aggregated Delay Costs (b) Delay Costs Per Transport Mode

Figure 3. Delay Costs for Different Numbers of Available Resources

We run our linear programming model on actual data to determine the number of resources that should be allocated to each vehicle for each period in order to minimize the total delay costs of all vehicles. This allocation is done with regard to the given workload and the number of available resources. The delay costs resulting from the optimal resource allocation are also indicated. The information about delay costs for different instances may be used to analyze and compare strategies at operational, tactical and strategic levels. At an operational level the model may be used to determine the number of operators to hire for the next day. Therefore, we run experiments on 20 different instances (20 days) to analyze the impacts of the number of available straddle carriers on delay costs. As expected, the number of instances with delays and the number of infeasible instances increase when the number of available resources decreases. However, in 95% of all instances, 16 out of 22 straddle carriers are sufficient to serve all vehicles without delays. Figure 3 illustrates our results. High delay costs for larger numbers of available resources result from instances that are infeasible for fewer resources. For infeasible instances, no delay costs may be calculated. We also see that if delays occur, truck delays make up a big part of the delay costs.

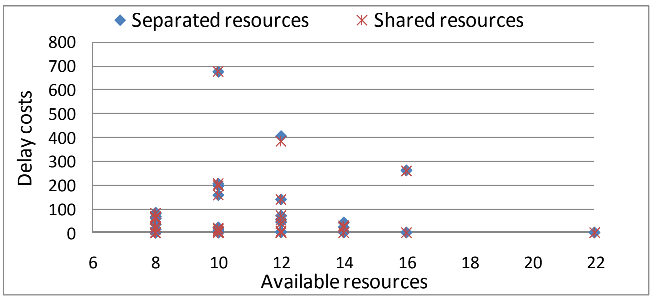


Figure 4. Comparison of Two Resource Allocation Strategies

At a tactical level, the optimization model may be used to compare the current resource allocation with alternative strategies. The current resource allocation does not allow sharing resources among different transport modes or among single barges and vessels. The alternative allocation strategy allows resources to be shared among all landside transport modes (trucks, trains and barges), but not among different vessels. Figure 4 presents the results of this comparison. We observe almost no difference for these two allocation strategies at the considered container terminal at GPMM. This is explained by the fact that little capacity remains unused for trucks and trains with the current strategy, as resources may be shared among different vehicles. Resources are allocated to single barges, which may lead to available resources being left unused. However, only one barge usually arrives per day and the unused capacity is very small. For all these reasons, the benefits from sharing resources among all landside transport modes are not significant.

4. A SIMULATION MODEL

Discrete event simulation is an important analysis technique to evaluate the performance of complex systems. In this study, we use simulation to evaluate resource allocation decisions at the container terminal, in particular the allocation proposed by the optimization model.

The simulation model representing one of the container terminals at GPM M is implemented in Arena 12. Containers, trucks, vessels, barges and trains are modeled as entities moving through the terminal. Straddle carriers and quay cranes are the system's resources. Import and export containers stocked at the yard are modeled as queues with infinite capacity. Since our main objective is to evaluate the resource allocation plan determined by the optimization model, we do not model the storage policies at the yard in detail. Our simulation model integrates the service of both inland transport means and vessels by straddle carriers.

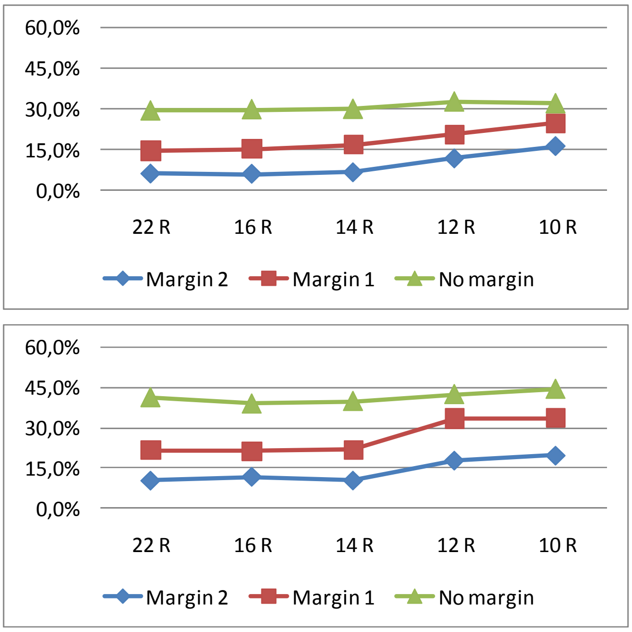
All straddle carriers are identical. They have the same speed and can transport only one container at a time. The optimization model calculates a resource allocation plan on an hourly basis. Therefore, straddle carriers are assigned to exactly one transport mode for each hour. Some external events, such as the arrivals of trains, vessels and barges, are known in advance. But, for the trucks, the mean arrival rate depends on the time of the day. We use a nonstationnary Poisson process to model truck arrivals, as this process is normally used to model time dependent arrival times.

Like the optimization model, the simulation model is composed of different and independent models representing different transport modes, and these models are related by the fact that they have to share a limited number of available resources. In addition, in the simulation model, they are connected by the flow of containers from landside transport means (trucks, trains, barges) via the yard to vessels (export containers) and vice versa (import containers). The yard is divided into two zones: one for import containers and the other for export containers. The modular structure enables us to represent the specific service characteristics of each transport mode. A short description is given below. A more detailed description is presented in Rodriguez-Verjan (2010). The truck model works as follows. When a truck arrives at the terminal gate, it may only enter the terminal if all required documents are available and if a parking space at the loading/unloading zone is free. If this is not the case, the truck has to wait. The loading/unloading zone has a given number of parking spaces. As soon as a truck parks at the loading/unloading zone, one of the straddle carriers allocated to trucks is scheduled to unload the container and transport it to the export zone of the yard. If the truck also loads an import container, a straddle carrier (not necessarily the same as for the unloading activity) picks up the import container from the import zone of the yard, transports it to the loading/unloading zone and loads it onto the truck. As soon as the truck is loaded, it exits the terminal. Its service time is registered to evaluate potential delays.

The procedures for vessels and barges are rather similar, even if vessels and barges differ with regard to service priority, volume and arrival rate. Therefore, only the vessel procedure is explained. When a vessel arrives at the terminal, it may enter if a position at the quay and quay cranes are available. As soon as the vessel has berthed, the unloading operation starts. Unloaded containers are placed in a buffer where they wait until a straddle carrier allocated to the vessel is scheduled to transport them to the import zone of the yard. Each vessel has its own buffer with a given container capacity. All import containers have to be unloaded before the loading operation may begin. When the unloading operation is finished, the loading operation starts. Straddle carriers allocated to that vessel pick up export containers at the export zone of the yard and transport them to the quay where they are loaded onto the vessel. As soon as the loading operation is finished, the vessel leaves the quay. Its service time is registered for later analysis.

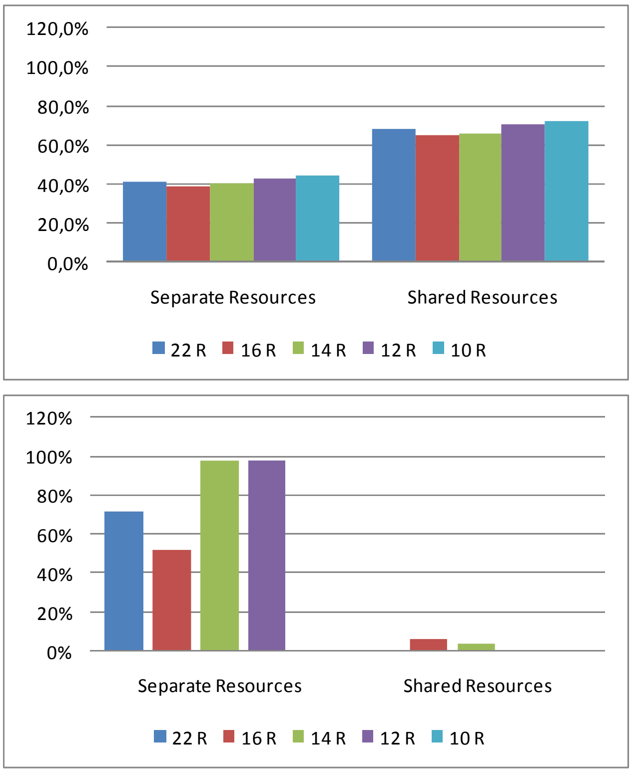
The train model is quite different from the previous models. As trains leave at predefined times, the goal is not to serve them as fast as possible but to assure that all containers have been unloaded/loaded prior to the train's departure time. The train station is modeled as a resource. When a train arrives at the terminal, the train uses this resource completely. The unloading operation begins and straddle carriers allocated to serve trains unload the containers and transport them to the export zone of the yard. When the unloading process is finished, the loading operation begins. As in the other models, the allocated straddle carriers pick up import containers at the yard and transport them to the station. The train leaves the terminal at the scheduled time even if the loading/unloading operation is not finished.

We analyze three different scenarios at different decision levels with our simulation model. At an operational level, we study the performance of a given resource allocation plan in a stochastic environment. At a tactical level, we analyze the impacts of different management policies for straddle carriers (e.g. shared or separated allocation between different transport modes). At a strategic level, we analyze the impacts of an increasing volume of containers transported by barges and trains that are more environmental friendly transport means. This is a pertinent scenario, since GPMM is investing in several projects to increase the number of containers transported by barges and trains.



(a) Low Variability (b) High Variability

Figure 5. Percentage of Delayed Trucks for Different Values of Variability of Straddle Carriers' Service Times



(a) Percentage of Delays for Trucks (b) Percentage of Delays for Vessels

Figure 6. Impacts of Separate and Shared Resource Allocation Policies

These three scenarios are tested for typical working days at the terminal. The following parameters represent these days: the expected number of containers lo load and unload per transport mode and the arrival rates for each transport mode. We also consider the physical configuration of the terminal which is the same for all scenarios.

For each scenario and each working day, several simulations are run. The input data of the simulations is the allocation plan and the variability of the straddle carriers' service times. In addition to this analysis, we also wanted to validate the consistency between the optimization model and the simulation model. Figure 5 shows that the resource allocation plan determined by the optimization model leads to good results in the simulation model. If a margin is added to the service time of containers, the allocation plan proposed by the optimization model is less sensitive to stochastic components of the system. If this margin is not included in the optimization, more delays may occur. We also observe that, in some cases with a low variability of the service times of straddle carriers, no additional resources have to be ordered. When the variability is larger, delays occur for the optimized allocation plan and additional resources are needed to reduce delays.

As the terminal does not use an appointment system and trucks are served in FIFO (First In First Out) order, delays at the beginning of the day have more impact on the global performance of the system than delays occurring at the end of the day. These results provide additional indications for allocation of priorities to be included in the optimization model. In Figure 6, we observe the impact on delays of vessels and trucks if straddle carriers are shared between inland transport modes and vessels. In this case, the allocation plan prescribes how many resources to use at the terminal, but does not specify the allocation to single transport modes. Vessels are served before trucks. In addition, the travel times of straddle carriers increase as they operate in a larger area. All this explains that delays for trucks may increase while delays for vessels decrease. With the current configuration and service priorities of the terminal, sharing transport resources between maritime and inland transport modes may harm inland service times and enhance vessel service times.

5. CONCLUSIONS AND PERSPECTIVES

This work proposes a mixed integer program to solve a resource allocation problem at container terminals. The aim is to serve all vessels during their time windows and minimize the delay of landside transport modes. This model was designed to represent a container terminal at GPMM. The model may be used by terminal operators to analyze problems at operational, tactical and strategic levels. Moreover, a discrete event simulation model representing the real terminal has been developed. This simulation tool allows terminal operators to evaluate the allocation plan before actual implementation, but also validates the optimization model. It should help to increase terminal operator's trust and visibility of possible effects of the allocation plan at an operational level.

This work can be pursued by adding new constraints to the models or adapting them to the situation at other container terminals operating with straddle carriers. It may for example be interesting to model tasks related to empty and freezer containers with their specific requirements. The dependency between quay cranes and handling resources or the limited berth place may also be modeled. Ou *et al.* (2010) show that scheduling arrivals may lead to decreasing waiting times. An interesting perspective is thus to adapt the model to be able to analyze the impacts of advancing or postponing the arrivals of trucks and barges on the number of necessary resources and delays. In the simulation model, the scheduling problem may be included. Currently, available resources are randomly allocated to available tasks, but an efficient dispatching strategy may lead to shorter travel times of straddle carriers and shorter service times of transport modes.

ACKNOWLEDGEMENTS

Part of this work has been conducted within the ESPRIT project, financed by the *Mission Transports Intelligents* of the DGITM (*Direction Générale des Infrastructures, des Transports et de la Mer*) of the MEEDDM (*Ministère de l'Ecologie, de l'Energie, du Développement Durable et de la Mer*).

The authors would like to thank Christophe Reynaud, from Marseille Gyptis International.

7. REFERENCES

Bish, E.K., Chen, F.Y., Leong, Y.T., Nelson, B.L., Ng, J.W.C. and Simchi-Levi, D. (2005). Dispatching vehicles in a mega container terminal. *OR Spectrum 27:* 491-506.

Chen, L., Bostel, N., Dejax, P., Cai, J. and Xi, L. (2007). A tabu search algorithm for the integrated scheduling problem of container handling systems in a maritime terminal. *European Journal of Operational Research, 181:* 40-58.

Egbelu, P.J. (1987). The use of non-simulation approaches in estimating vehicle requirements in an automated guided based transport system. *Material Flow, 4:* 17-32.

European Sea Ports Organization (ESPO) (2004). Factual report on the European port sector.

Froyland, G., Koch, T., Megow, N., Duane, E. and Wren, H. (2006). Optimizing the landside operation of a container terminal, *OR Spectrum, 30: 53 -75.*

Gambardella, L.M., Rizzoli, A.E. and Zaffalon, M. (1998). Simulation and planning of an intermodal container terminal. Special Issue in Harbour and Maritime Simulation, *Simulation, 71, 2: 107-116.*

Gambardella, L.M., Mastrolilli, M., Rizzoli, A.E. and Zaffalon, M. (2001). An optimization methodology for intermodal terminal management. *Journal of Intelligent Manufacturing, 12:* 521-534.

Hartmann, S. (2004). A general framework for scheduling Equipment and Manpower at Container Terminals. *OR Spectrum, 26:* 51-74.

Huynh, N. and Walton, C.M. (2008). Robust scheduling of truck arrival at marine container terminals. *Journal of Transportation Engineering, 134:* 347-353.

Lau, H.Y.K. and Zhao, Y. (2008). Integrated scheduling of handling equipment at automated container terminals. *International Journal of production economics, 112:* 665-682.

Namboothiri, R. (2006). Planning container drayage operations at congested seaports. *PhD thesis.* Georgia Institute of Technology.

Ou, J., Hsu, V. and Li, C.L. (2010). Scheduling truck arrivals at an air cargo terminal. *Production and Operations Management, 19:* 83-97.

Rodriguez-Verjan. G.L. and Dauzère-Pérés (2010). Optimisation des opérations en logistique protuaire. *Master thesis.* Ecole des Mines de Saint Etienne.

Sanchoy, K.D. and Spasovic, L. (2003). Scheduling material handling vehicles in a container terminal, *Production Planning & Control 14*: 623-633.

Stahlbock, R. and Voss, S. (2008). Operation research at container terminals: a literature update. *OR Spectrum, 30:* 1-52.

Steenken, D., Voss, S. and Stahlbock, R. (2004). Container terminal operation and operations research - A classification and literature review. *OR Spectrum, 26*: 3-49.

Vis, I.F.A., Koster, R. and Salvelsbergh, M.W.P. (2005). Minimum vehicle fleet size under time-window constraints at a container terminal. *Transportation Science, 39:* 249-260.

Zehendner, E., Dauzère-Pérés, S., Feillet, D., and Absi, N. (2010). Optimising the resource allocation at the Port of Marseilles Container Terminal. *Master thesis.* Grenoble INP - Génie Industriel.

Session D3: Seaport and Transportation 4

·Day1: Sep. 15 (Wed.)

·Time: 16:40 - 18:00

·Chair: Stefan Voß

·Room: Iris, 4F

EMPTY CONTAINER MANAGEMENT – STATE OF THE ART

Robert Stahlbock[1] and Stefan Voß[1]

[1]University of Hamburg
Germany
e-mail: stahlbock@econ.uni-hamburg.de
stefan.voss@uni-hamburg.de

Abstract: Billions of dollars are spent each year to deal with inefficiencies caused by repositioning of empty containers resulting from an imbalanced cargo flow between the continents. Therefore, the logistics of empty containers is receiving a lot of attention from both practitioners and researchers. They are focused on various problems such as the optimization of the container flow for providing proper repositioning, the provision of related data, the forecasting of demand and supply on a global, a regional, and a local level, and for different time horizons, but also the overcoming of the imbalance from an economical and ecological point of view. Cargoes which have helped to reduce the trade imbalance include scrap steel, plastic and waste paper. However, additional ideas and concepts are necessary to improve the situation. We consider the problem of empty container transport. We briefly describe the global chain of maritime container transport and related costs for different actors. We provide an analysis of data emphasizing the imbalance of loaded and empty container transportation between various origins and destinations. We provide a review of literature regarding empty container management and discuss some approaches for reducing empty container transportation and costs, e.g., by using foldable containers. We consider ecological issues related to reverse logistics (e.g., scrap, waste paper/recycle). We propose the idea of pooling (i.e., interchanging containers between shipping companies, owners, etc., on different scales). We provide basic ideas regarding forecasting approaches as well as appropriate network design and multi-agent approaches. Finally, we discuss entry and exit of containers with respect to the transport chain.

OPERATING YARD CRANES OVER STORAGE BLOCKS TO MINIMIZE THE HANDLING TIME IN A CONTAINER YARD

Shiwoo Lee[1*] and Weerakit Denkaew[2]

[1]Division of Mechanical and Industrial Engineering,
Pohang University of Science and Technology,
Pohang, Korea
e-mail: cameolee@postech.ac.kr

[2]School of Mechanical and Industrial, Manufacturing Engineering,
Oregon State University,
Corvallis, OR, USA

Abstract: In container terminals, the efficiency of loading operations in a container yard highly depends on the productivity of yard cranes. In loading operation, yard cranes have to pick-up the containers which are stacked in the yard-bays to satisfy the work schedule requirement of quay crane, the equipment for loading and unloading containers to and from a container ship. This paper focuses on operating multiple yard cranes among storage blocks, allowing multiple yard cranes over a block of yard-bays. The objective of the pick-up scheduling problem is to minimize the total container handling time, which consists of the traveling distance and the setup time of yard cranes in a container yard. A mixed integer programming model has been proposed. Since it is very difficult to obtain the optimal solution, a genetic algorithm has been proposed to solve the pick-up scheduling problem for operating multiple yard cranes over storage blocks. Computational results show that the proposed GA is an effective method.

1. INTRODUCTION

In international sea freight transportation, the use of containers for general cargos has significantly increased over the past decades. Therefore, container terminals take an important role in the transportation network by providing services to customers worldwide. The operations in container terminals have to be improved to satisfy the increasing demand of the customers. The total spending for container management in containerized transportation industry was estimated to be in the region of \$100 billion to \$110 billion per year. From this total expense, approximately 11% to 15%, or \$14 billion to \$16.5 billion annually, was believed to be associated with inefficiencies of container terminal operations (Behenna, 2001).

Planning container terminal operations consists of berth scheduling, quay crane (QC) scheduling and load sequencing (Jung and Kim, 2006). Many researchers have focused on the QC scheduling and load sequencing with the main objective to increase the container terminal throughput. We focus on load sequencing of yard side equipment (YSE), such as yard cranes, transfer cranes, and straddle carriers. The load sequencing generally consists of container loading/unloading process and dispatching process of YSE. The problem of load sequencing consists of two subproblems. The first subproblem is to determine a sequence of yard-bays that each YSE visits. The second subproblem is to determine the number of containers to be picked-up at each yard bay. The objective of the load sequencing is to minimize the container handling time of the YSE. This problem is also known as the "pick-up scheduling problem". Figure 1 shows a typical layout of a container terminal.

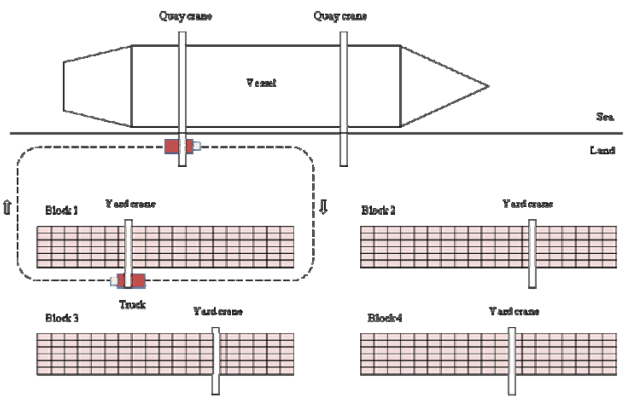


Figure 1. A Typical Layout of Container Terminal

Generally, when a container vessel arrives at a container terminal, it will be served by multiple quay cranes. In an unloading operation, imported containers are discharged from the vessel and loaded onto trucks or straddle carriers (see Figure 1.5) by quay cranes. The containers are then unloaded quickly to various locations in the container yard. On the other hand, in loading operation, exported containers in the container yard are loaded onto trucks by container handling equipments, e.g. yard cranes or straddle carriers. Then the containers are transported to loading areas by other YSE, for instance trucks or straddle carries. Thereafter, the containers are loaded onto a container ship or a vessel by quay cranes.

We have focused on the operation of multiple yard cranes during loading operation in a container yard. Yard crane was derived from industrial overhead cranes. There are two different types of yard cranes which are rubber-tired gantry crane (RTGC) and rail-mounted gantry crane (RMGC). The first one is frequently referred to transtainer and runs on heavy-duty pneumatic-tired wheels and the latter one runs on steel wheels over fixed rails. The containers within the storage area are normally arranged in long rows parallel to the wharf, with about 30 TEU slots per row, which is thus approximately 180 m long. There are usually five or six rows of containers per block, and the truck lane. At the ends of the blocks, a roadway space is provided between adjacent blocks. The wheels of the RTGC can turn through 90 degrees so that it can move from one storage block to another as required to meet operational needs while RGMC can only move along the fixed rails.

At storage blocks in the container yard, yard cranes operate for all containers handling works, which include lifting containers for trucks, storing them at storage locations, picking up containers from the storage locations and putting them on the trucks. They play a very important role in the container yard and are usually the bottleneck in the container handling process (Zhang, Wan, Liu, and Linn, 2002). Therefore, in order to achieve a good productivity and a high efficiency, the deployment of yard cranes should be planned carefully. Problems of scheduling vehicles and material handling equipment have been arisen frequently in logistic systems and have also been extensively studied under various different settings (Bramel and Simchi-Levi, 1997 as cited in Ng, 2005). However, the results reported in a vast majority of research literatures were not directly applicable to a container terminal due to its unique characteristics (Ng, 2005).

A particular case of the pick-up scheduling problem for a single yard crane was studied by Kim and Kim (2002). They studied the problem of one yard crane operated in a single block in a container yard. Their objective was to minimize the container handling time, which consists of the total traveling distance of the yard crane and the setup time of the yard crane, by determining the sequence of yard-bays to be visited and the number of containers to be picked-up at each yard-bay. Kozan and Preston (1999) studied the scheduling of transfer operation of containers without considering the actual sequencing issue. The research of a single yard crane operation was also studied by Kim and Kim (1999a), (1999c) and Narasimhan and Palekar (2002).

The operation of multiple yard cranes is also considered as one of the pick-up scheduling problems. This operation was studied by Jung and Kim (2006). They proposed a method to schedule loading operations for multiple yard cranes which operate over a storage block. In accordance with multiple yard cranes operate in a storage block, the interference between adjacent yard cranes is significantly important to be considered.

Recently, Lee, Zhi, and Meng (2007) studied the scheduling problem of multiple yard cranes for loading containers over storage blocks. However, they limited the accessibility of each yard crane by assuming that the yard crane was not allowed to travel between two blocks during the loading process.

Linn and Zhang (2003) studied yard cranes deployment problem in each time period. In their research, multiple yard cranes were assigned to work over multiple storage blocks. However, they did not consider either the scheduling of detailed movements for each yard crane or particular tasks for each yard crane during the loading operation. The problem of scheduling multiple yard cranes in a container yard to minimize the total jobs' waiting time was also studied by Ng (2005). The author addressed the scheduling problem of yard cranes by considering the yard cranes interference among adjacent blocks. Both of the specific tasks during the loading operation and the detailed movement of each yard crane were also not considered in the scheduling problem.

2. PROBLEM STATEMENT

In a container terminal, a container yard is divided into rectangular regions called *blocks*. Each block typically consists of 20-30 *yard-bays*, which store a number of containers. A container type is defined as a collection of containers which have the same size and the same delivered destination. It is assumed that each yard-bay can store only one type of containers. Examples of a work schedule and a yard map are given as shown in Table 1 and Figures 2, respectively. The work schedule shows the number of containers of each type to be picked-up sequentially for quay cranes. The yard map shows the distribution of containers of each container type in the container yard. A "*sub-tour*" denotes a sequence of containers of each container type to be picked-up during loading operations.

The objective of the pick-up scheduling problem is to minimize the total container handling time, which consists of the traveling distance between visiting yard-bays, and the setup time of yard cranes on the yard-bay. The traveling distance of a yard crane is dependent on the visiting sequence of yard-bays. The setup time of yard crane occurs whenever the yard crane moves from one yard-bay to another. Even though a yard crane is usually assigned to a block, it might be able to move to operate in the other blocks. In this research, it is assumed that multiple yard cranes that have the same features are operating in a container yard during the loading operation. It is also assumed that there are a sufficient amount of trucks for delivering containers from container yard to loading zones, which means the waiting of the trucks can be ignored.

Table 1. An Example of A Work Schedule

Sequence (sub-tour index)	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Container type	D	C	A	D	E	D	G	A	B	C	B	H	H	G	F	E
Quantity	27	17	32	57	27	3	8	5	39	22	47	81	22	5	77	6

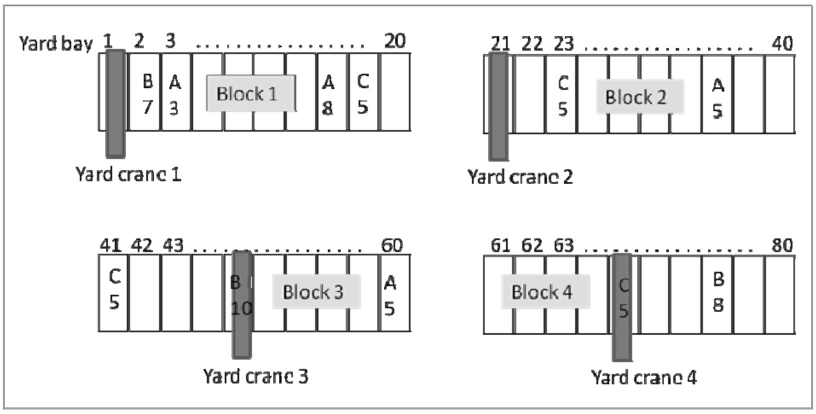


Figure 2. An Example of A Yard Map

When loading a container from the container yard to a ship, a yard crane moves back and forth along the specific block to the intended container which is stacked in a yard bay, then its hoist picks up the container and loads it onto a truck. The container has to be delivered to quay crane in the order specified in the work schedule. Therefore, all yard cranes must operate in the order prescribed by the work schedule of quay cranes. Afterwards, the truck transports the container to a corresponding quay crane, which picks up the container and loads it onto a ship. The efficiency of loading operations significantly depends on the loading schedule of the exported containers to be loaded on the ship. Thus, the efficient container handling at container terminal is important in reducing container transportation costs and keeping shipping schedules, since it has to satisfy the work schedule required by quay crane.

The main goal of this paper is to optimize the operation of yard cranes in a container terminal. The first objective is to propose a mathematical model of the pick-up scheduling problem for a single yard crane operating over a storage block. The second objective is to develop a mathematical model of the pick-up scheduling problem for multiple yard cranes operating over a number of storage blocks. The third objective is to propose a heuristic algorithm, i.e., genetic algorithm (GA) to effectively and efficiently solve the medium and large size of the pick-up scheduling problem under consideration.

3. MATHEMATICAL MODEL

A mathematical model for the operation of a single yard crane, originally proposed by Kim and Kim (2002), has been extended to a mathematical model representing the operation of multiple yard cranes over storage blocks. In addition, we have removed the set operators which were in the model by Kim and Kim (2002). This section provides a mathematical formulation for the pick-up scheduling problem of operating multiple yard cranes over storage blocks.

The following assumptions have been used and the additional assumptions are given as follows:

- There is only one type of containers stacked in a yard bay.
- A yard crane is working in a single block during loading operation.
- A yard crane is initially located at the first yard bay of the block.
- Multiple yard cranes of the same features are working together over multiple blocks during the loading operation.

- A yard crane is assigned to each storage block. In other word, the number of yard crane is equal to the number of block.
- Each yard crane is initially located at the first yard bay of each block.
- When a yard crane moves to a yard-bay within or outside of the current storage block, the shortest path will be chosen.

The penalty rule applying when a yard crane moves to a yard-bay outside of the current storage block is described in the following paragraphs.

Notations:

i,j	yard-bay index
d(i,j)	distance between i and j
n	the number of yard-bays per a block
$B = [\ ]^{pxq}$	the matrix of blocks (p = the number of row of blocks, q = the number of column of blocks, p*q = the number of blocks)
(r, c)	the coordination of the block (r = the index of row, c = the index of column)
(r_1, c_1)	the coordination of the block in which yard-bay i exists (starting location)
(r_2, c_2)	the coordination of the block in which yard-bay j exists (destination)

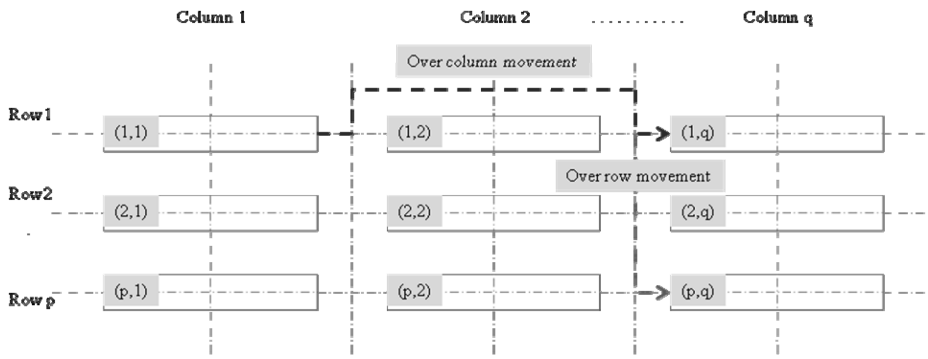


Figure 3. The Penalty Rule for A Yard Crane Movement

Figure 3 shows the possible yard crane movements which are "over column movement" and "over row movement". We want to know the traveling distance from yard-bay i to yard-bay j. Therefore, we need to know which blocks where yard-bay i and j exist, so that we can calculate the penalty of the movement between the blocks. A procedure for calculating the traveling distance between two yard-bays is described by the following steps.

Step 1: Indicate the coordination of the block in which yard-bay i(j) exists as follow:

- Determine the row of the block in which yard-bay i(j) exists ("\" indicates *integer division* operator)
  $\mathbf{r} = [i+(n*q-1)] \backslash n*q$
- Determine the column of the block in which yard-bay i(j) exists
  If $([i+(q-1)] \backslash n)$ *mod* $q = 0$ then $\mathbf{c} = b + q$
  Else $\mathbf{c} = b$

Step 2: Apply the procedure in Step 1 to the index j. Then we will have the coordination of the blocks in which yard-bay i and j exist respectively.

Step 3: After we get the coordination (r_1, c_1) and (r_2, c_2), the penalty rules of yard crane movement can be determined as follows.

If $|c_1 - c_2| <= 1$ (adjacent column) then Use the procedure for adjacent columns in the following.
Else (i.e., $|c_1 - c_2| > 1$) $d(i,j) = |c_1 - c_2 - 1|\alpha + |r_1 - r_2 - 1|\beta$

, where α = constant when yard crane move over a column, β = constant when yard crane move over a row.

Note that when yard cranes move over columns, the different penalty rules can be applied. However, when yard cranes move between adjacent columns, the fewer penalties are applied.

Figure 3.2 shows the potential movements of a yard crane between two blocks in the adjacent columns. If two blocks are adjacent to each other and align longitudinally, for example block 1 and block 2, a yard crane can move from one block to the other one easily without any turning movement (see Path1 in Figure 3.2). However, a penalty rule is applied since moving a yard crane between two blocks might cause the delay for other operations. In the other case, if a yard crane moves from a block to the other one on a different row, for example, between block 1 and block 4, more penalties are applied to this movement. The yard crane must first move out of the current block, make a 90º turn and move to the other row, then make a 90º turn again and enter the other block (see Paths 2 and 3 in Figure 4). This movement occupies a large amount of traffic space for an extended time period, which can obstructs the traffic on the roads between blocks and also delays other terminal operations.

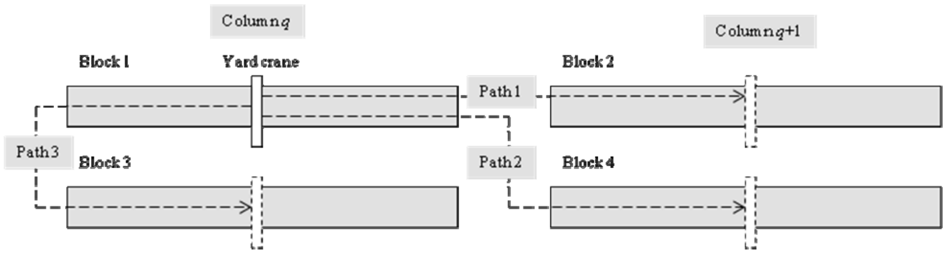


Figure 4. Potential Movements of Yard Crane Between Blocks in The Adjacent Columns

The notations used in the model are given as follows:

i, j	The indices for yard bays
v	The index for yard cranes
m	The number of sub-tour that constitute a complete tour of a yard crane
Ω	The set of all yard-bays
Φ	The set of empty yard-bays (the yard-bay that does not have any container)
l	The number of container types
V	The number of yard cranes
$S(h)$	The set of sub-tour numbers in which the containers of type h exist
$B_v(h)$	The set of yard-bay numbers in which the containers of type h exist, and are attended by yard crane v
c_{hj}	The initial number of container of type h stack at yard-bay j
r_t^h	The number of containers of type h to pick-up during sub-tour t
g_t	Container type number of sub-tour t
d_{ij}	Travel distance between yard-bay i and j
e_{ij}	0, if $i = j$, 1, otherwise,
t	Sub-tour number, $t = 0, 1, \ldots, m$, $m+1$, where $t = 0$ and $m+1$, respectively, correspond to the initial and the final locations of the yard crane
S_v	The initial location of the yard crane v
F_v	The final location of the yard crane v
T_s	The setup time of yard crane for each visit to a yard bay
T_d	The traveling distance of yard crane per the distance of a yard bay length

Decision variables:

y_{vij}^t	1, if yard crane v moves from yard-bay i to yard-bay j after completing sub-tour t 0, otherwise
z_{vij}^t	1, if yard crane v moves from yard bay i to yard bay j during sub-tour t 0, otherwise
x_{vj}^{ht}	The number of container of group h picked up at yard bay j during sub-tour t by yard crane v

A mixed integer programming model for multiple yard cranes over the multiple storage blocks has been presented as follows:

$$\text{Minimize} \sum_{t=0}^{m}\sum_{v=1}^{V}\sum_{i,j\in\Omega}(T_s e_{ij}+T_d d_{ij})y_{vij}^{t}+\sum_{t=1}^{m}\sum_{v=1}^{V}\sum_{i,j\in\Omega}(T_s+T_d d_{ij})z_{vij}^{t} \quad (1)$$

$$\text{Subject to} \sum_{v}^{V}\sum_{j\in B_v(g_1)} y_{vS_v j}^{0}=1, \quad (2)$$

$$\sum_{v}^{V}\sum_{j\in B_v(g_m)} y_{vjF_v}^{m}=1, \quad (3)$$

$$\left(\sum_{v}^{V}\sum_{j\in\Omega}y_{vji}^{t-1}+\sum_{v}^{V}\sum_{k\in\Omega}z_{vki}^{t}\right)-\left(\sum_{v}^{V}\sum_{j\in\Omega}y_{vij}^{t}+\sum_{v}^{V}\sum_{k\in\Omega}z_{vik}^{t}\right)\le 1, \qquad i\in\Omega, t=1,2,\ldots,m, \quad (4)$$

$$\sum_{i\in\Omega}\sum_{j\in\Phi}\sum_{v}^{V} y_{vij}^{t}=0, \qquad t=0,1,\ldots,m \quad (5)$$

$$\sum_{i\in\Omega}\sum_{j\in\Phi}\sum_{v}^{V} z_{vij}^{t}=0, \qquad t=1,2,\ldots,m \quad (6)$$

$$x_{vj}^{ht}\le M\left(\sum_{v}^{V}\sum_{k\in\Omega}z_{vkj}^{t}+\sum_{v}^{V}\sum_{i\in\Omega}y_{vij}^{t-1}\right), \qquad j\in\Omega, h=1,2,\ldots,l, v=1,2,\ldots,V, t=1,2,\ldots,m, \quad (7)$$

$$\sum_{j\in\Omega}\sum_{v}^{V}x_{vj}^{ht}=r_t^h, \qquad h=1,2,\ldots,l,\ t=1,2,\ldots,m, \quad (8)$$

$$\sum_{t=1}^{m}\sum_{v}^{V}x_{vj}^{ht}=c_{ht}, \qquad h=1,2,\ldots,l,\ j\in\Omega, \quad (9)$$

$$y_{vij}^{t}\in\{0,1\}, \qquad i,j\in\Omega, v=1,2,\ldots,V, t=1,2,\ldots,m, \quad (10)$$

$$z_{vij}^{t}\in\{0,1\}, \qquad i,j\in\Omega, v=1,2,\ldots,V, t=1,2,\ldots,m, \quad (11)$$

$$x_{vj}^{ht}\ge 0, \qquad j\in\Omega \quad h=1,2,\ldots,l, \quad v=1,2,\ldots,V, t=1,2,\ldots,m \quad (12)$$

The objective function (1) is to minimize the total container handling time, which depends on the total travel distances of every yard crane and the setup time which occurs whenever yard crane travel between yard-bays. Constraints (2), (3) represent the flow conservation of yard cranes at the initial and the final locations. Constraint (4) represents the flow conservation at the other locations. Constraint (5), (6) ensure that yard crane will not move to yard-bay that does not have any container. Constraint (7) implies that the containers at a yard bay can be picked up only when yard cranes operate at the yard bay. Constraint (8) ensures that the number of containers picked up in a sub tour must be equal to the number of containers requested by the work schedule. Constraint (9) ensures that the number of containers picked up during the whole tour must be equal to initial number of containers at each yard bay for every container type. Constraint (10), (11) ensures that y_{vij}^{t}, and z_{vij}^{t} can be 0 or 1. Constraint (12) ensures a nonnegative integer.

4. A GENETIC ALGORITHM AND ITS COMPUTATIONAL RESULTS

The pick-up scheduling problem is difficult to be solved optimally in a reasonable amount of time even though the problems are small. Apparently, the mathematical model might not be able to solve a large sized problem within a reasonable computational time. The GA has been applied successfully in many combinatorial optimization problems. However, it does not guarantee the optimality according to its stochastic nature, but it can rapidly locate a good solution or near-optimal solution in significantly less time. The proposed GA was developed to efficiently solve the medium and the large sized problems. The detail of the proposed GA is left out for the brevity.

Considerable experiments of various parameter combinations including the probability of crossover (CR), the probability of mutation (MR), and the offspring choosing strategy (OCS) were performed to obtain the most efficient parameter combination to make an objective value converged to a stabilized value. A test problem is exemplified in the following. There were 8 types of containers distributed in 10 storage blocks. Each storage block contains 17 yard-bays.

Figure 5 shows the given yard map. A work schedule illustrated in Table 2 is also randomly created for this particular test problem.

Table 2. Work Schedule of The Test Problem

Sequence (sub-tour index)	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Container type	D	C	A	D	E	D	G	A	B	C	B	H	H	G	F	E
Quantity	27	17	32	57	27	3	8	5	39	22	47	81	22	5	77	6

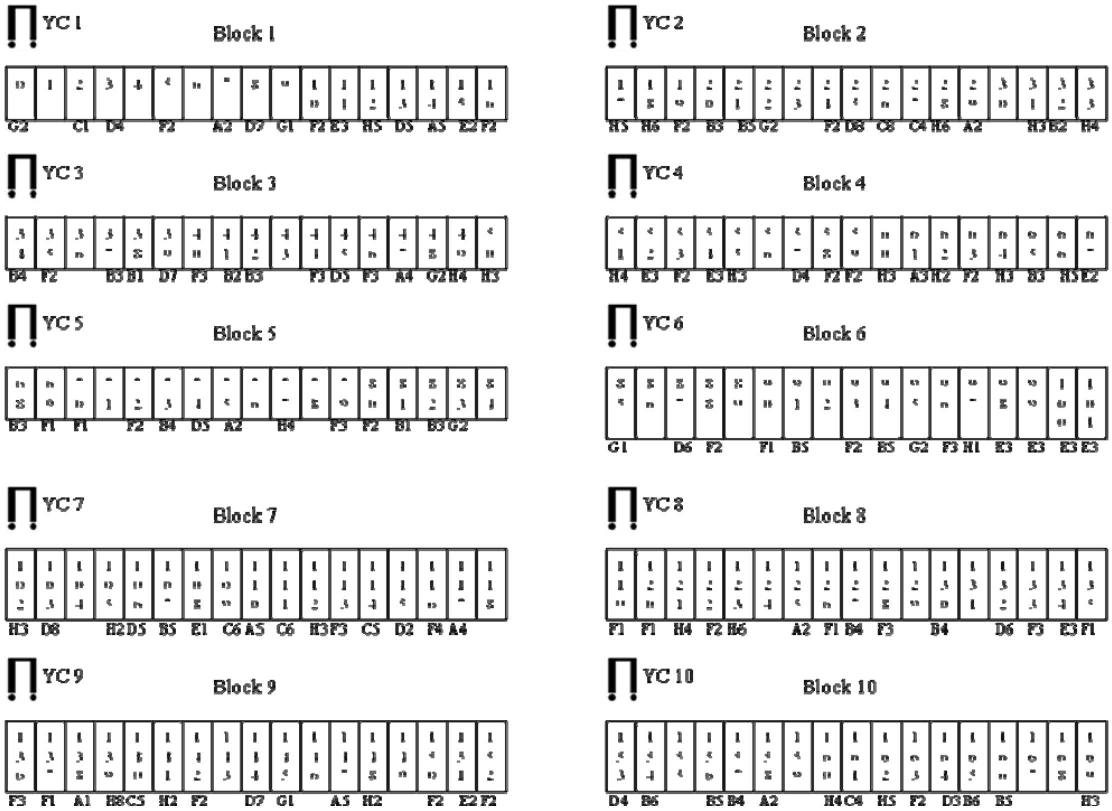


Figure 5. Yard Map of The Test Problem

Each yard crane is initially located at the first yard-bay of each block as assumed. The traveling distance of the yard crane between adjacent yard-bays within a block is set to be 1 unit. Whenever the yard crane moves from its block to another block, the penalty of the movement is applied.

For the proposed GA, the population size and the number of generations were set to 50, and 5000, respectively. However, the algorithm stopped whenever the solution value was not improved within 500 generations. The tested values of the probability of crossover were 0.7, 0.8, and 0.9. The tested values of the probability of mutation were 0.05, 0.15, and 0.25. The offspring choosing strategies which rejected duplicated offsprings (OCS1) and accepted duplicated offsprings (OCS2) were tested. Each parameter combination was run for 5 times. The performance of each parameter combination was evaluated by considering average solution value, range of the solution value, and the variance of the solution value. The parameter combination which provided the most consistent and most stable solution value was chosen.

A statistical program was used to statistically analyze the factors which had effects to the performance of the proposed GA. The parameter combination using crossover rate 0.8, mutation rate 0.05 with OS1 is used in the proposed GA to demonstrate the effectiveness of the proposed GA, since it is robust to solve the medium and large sized problems.

Table 5.2 Solution Values from Various Parameters Combination

OCS	CR	MR	Run 1	Run 2	Run 3	Run 4	Run 5	Average	Range	Variance
OCS1	0.7	0.05	1608	1598	1602	1593	1596	1599.4	15	5.81
		0.15	1567	1575	1589	1570	1565	1573.2	24	9.60
		0.25	1560	1586	1593	1587	1597	1584.6	37	14.47
	0.8	0.05	1562	1563	1555	1567	1563	1562.0	12	4.36
		0.15	1586	1577	1584	1573	1588	1581.6	15	6.35
		0.25	1579	1563	1571	1552	1557	1564.4	27	10.81
	0.9	0.05	1572	1582	1573	1596	1593	1583.2	24	11.08
		0.15	1568	1592	1589	1562	1588	1579.8	30	13.75
		0.25	1598	1596	1584	1562	1575	1583.0	36	15.00
OCS2	0.7	0.05	1605	1607	1598	1588	1591	1597.8	19	8.35
		0.15	1566	1594	1590	1602	1585	1587.4	36	13.48
		0.25	1561	1591	1589	1602	1575	1583.6	41	15.87
	0.8	0.05	1592	1592	1579	1588	1573	1584.8	19	8.47
		0.15	1587	1564	1561	1593	1593	1579.6	32	15.84
		0.25	1568	1556	1566	1575	1565	1566.0	19	6.82
	0.9	0.05	1590	1567	1603	1601	1582	1588.6	36	14.77
		0.15	1576	1558	1562	1589	1573	1571.6	31	12.26
		0.25	1571	1544	1590	1566	1588	1571.8	46	18.71

5. CONCLUSIONS

The MIP model of operating a single yard crane over a block has been extended to the MIP model of operating multiple yard cranes over storage blocks. A genetic algorithm for solving the pick-up scheduling problem has been proposed. The computational results from the experiments show that the proposed GA is an effective method for solving the pick-up scheduling problem of operating multiple yard cranes over storage blocks during a loading operation in a container yard.

6. REFERENCES

Behenna P. (2001), "Container Asset Visibility: The Key to Reducing Ocean Carrier Operating Costs", *International Asset Systems*, 1-9.

Jung S.H., and Kim K.H. (2006), "Load scheduling for multiple quay cranes in port container terminals", *J Intell Manuf*, (17), 479-492.

Kim K.Y., and Kim K.H. (1999a), "A routing algorithm for a single straddle carrier to load export containers onto a containership", Int*. J. Production Economics 59*, 425-433.

Kim K.H., and Kim K.Y. (1999b), "Routing straddle carriers for the loading operation of containers using a beam search algorithm", *Computers & Industrial Engineering*, 36 , 109-136.

Kim K.H., and Kim K.Y. (1999c), "An optimal routing algorithm for a transfer crane in port container terminals", *Transportation Science*, 33 (1), 17-33.

Kim K.Y., and Kim K.H. (2002), "Heuristic algorithms for routing yard-side equipment for minimizing loading times in container terminals", *Naval Research Logistics*, 50, 498-514.

Kozan E., and Preston P. (1999), "Genetic algorithms to schedule container transfers at multimodal terminals", *International Transactions in Operational Research*, Res. 6 331-329.

Lee D.H., Zhi C., and Meng Q. (2007), "Scheduling of two-transtainer systems for loading outbound containers in port container terminals with simulated annealing algorithm", *Int. J. Production Economics*, 107, 115-124.

Linn J.R., and Zhang C.Q. (2003), "A heuristic for dynamic yard crane deployment in a container terminal", *IIE Transactions*, 35, 161-174.

Narasimhan A., and Palekar U.S. (2002), "Analysis and algorithms for the transtainer routing problem in container port operations", *Transportation Science*, Vol.36, No.1, 63-78.

Ng W.C. (2005), "Crane scheduling in container yards with inter-crane interference", *European Journal of Operational Research*, 164, 64-78.

Zhang C., Wan Y.W., Liu J., and Linn R.J. (2002). "Dynamic crane deployment in container storage yards", *Transportation Research Part B*, 36, 537-555.

SCHEDULING RESHUFFLE OPERATIONS AND DOUBLE CYCLES ON A SINGLE QUAY CRANE

Christian Bierwirth[1] and Frank Meisel[1*]

[1]Martin-Luther-University
Germany
e-mail: christian.bierwirth@wiwi.uni-halle.de
frank.meisel@wiwi.uni-halle.de

Abstract: Fast handling of vessels is one of the most important goals in container terminal operations planning. Recent studies indicate that the handling of vessels is significantly accelerated by quay crane double cycling. In our paper, we show that the service process can be further accelerated by changing the treatment of so-called reshuffle containers. Reshuffle containers have to be removed from their position in the vessel only to gain access to containers stacked below them. Our approach enables to reposition reshuffle containers directly within the bay of a vessel, referred to as internal reshuffles, instead of temporarily unloading them. We develop a MIP model and provide a fast heuristic for building crane schedules. Computational tests prove that the consideration of internal reshuffles leads to a further shortening of vessel handling times compared to a sole application of crane double cycling.

DEVELOPMENT OF E-MARKETPLACE SYSTEM FOR IMPROVING THE MARKET STRUCTURE OF KOREAN OVERLAND TRANSPORTATION

Hyung Rim Choi[1], Hyun Soo Kim[2], Soon Goo Hong[3], Min Je Cho[4*], Ki Nam Choi[5], Chang Hyun Park[6], and Jeong Soo Ha[7]

Department of Management Information Systems
Dong-A University
Bumin-dong, Seo-gu
Busan, Korea
{[1]hrchoi, [2]hskim, [3]shong}@dau.ac.kr, [4]mjcho78@gmail.com
{[5]dreizehn, [6]archehyun}@naver.com, [7]charismaya@gmail.com

Abstract: The overland transportation market consists of three major participants; cargo owners (shippers) to make orders to transport, car owners (trucking companies) to delivery cargoes and agencies to connect the two aforementioned players. However, in case of insufficiency in sharing the information between transportation demanders and suppliers, agencies or other types of mediators participating in the transportation market with multilevel marketing, high brokerage in the total transportation margin causes car owners' weak profitability. On the other hand, there is another problem that freight charges are fixed by shippers but costs grow steadily because of increasing fuel and others costs. Therefore, car owners need to find out higher-priced cargos in the market.

To solve these problems, a new type of agency is needed. Thus, this study aims to develop an online e-marketplace model to connect each other freely between shippers, car owners and the other related parties without high brokerage or commission.

We studied together with developing an e-marketplace systems and its successful operating strategies. The new e-marketplace will contribute to reduce the brokerage between cargo and car owners. In addition also its implementation strategies will activate users to participate into the system.

1. INTRODUCTION

Among the pending issues surrounding the freight transportation market, in particular, problems of the unnecessary multi-level structures, high empty car ratios and supply-demand imbalance etc are causing many social problems and online freight transportation transactions are recognized as one of the method for solving these problems. Although online freight transportation transaction services are being provided now, the services are not frequently utilized. Moreover, studies of the online services in the freight transportation market are extremely rare. In this respect, this paper is intended to problems in online freight transportation transactions and success factor for activating online freight transportation transactions.

To this end, in this study, the present state of the freight transportation market in Korea and problems in the market were examined first through literature study. Then, the present state of online transportation transactions in foreign countries and the success factors of existing e-Marketplaces were summarized. In addition, the perception of online freight transportation transactions and related problems were grasped through expert interviews in order to present measures to activate e-Marketplaces in the freight transportation market.

2. PRESENT STATE OF THE FREIGHT TRANSPORTATION MARKET

2.1. Overview of The Freight Transportation Market and Its Importance

In the freight transportation market, truck transport businesses generally refer to truck transportation businesses, truck transportation brokerage businesses and truck transportation franchise businesses. And truck transportation businesses refer to businesses to transport freight at a cost using trucks as requested by others. Truck transportation brokerage businesses refer to mediating freight transportation contracts at a cost as requested by others or transporting properties under one's own title using the freight transportation measures of those who operate truck transportation businesses or truck transportation franchise businesses. And truck franchise businesses refer to transporting freight at a cost using one's own trucks or having affiliated truck transportation franchisees transport the freight. These truck transportation related businesses in Korea are divided and summarized in Figure 1.

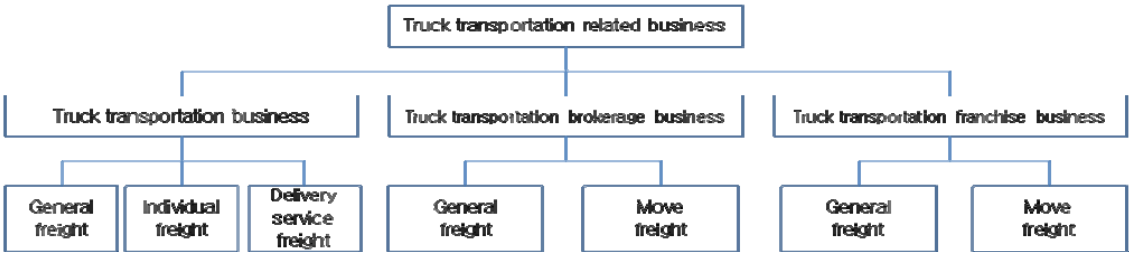


Figure 1. Truck Transportation Related Businesses

Although these truck transportation related businesses are service businesses, they can be said to be an important industry that serves the role of arteries of economic flows. In particular, the weights of logistical costs in domestic businesses are higher compared to advanced countries and this becomes a major factor to reduce domestic businesses' competitiveness. Domestic logistical costs reach 106 trillion won as of 2006 and this accounts for 12.5% of our country's GDP. Of them, land transportation through public roads occupies the highest percentage at 76.2%(as of 2008) and thus it can be seen that the effect of freight transportation on domestic economy is large. Therefore, the improvement and efficient management of the freight transportation market can be said to be very important. Recently, the government has been implementing diverse policies to solve chronic problems in the freight transportation market. Among these policies, along with the recent development of information technologies, the government is politically supporting and implementing the introduction and utilization of online freight transportation markets, that is, e-Marketplaces as a way to solve the chronic problems in the freight transportation market.

2.2. Present State of The Freight Transportation Market and Problems In The Market

Freight transportation business entry systems existed until 1999 as the registration system and the license system and the businesses were quantitatively controlled by the knack of determining supply criteria and guidelines for licence issues. However, in the situation where unemployment increased due to the IMF foreign exchange crisis, freight transportation business reporting systems were changed from permission systems to registration systems in July 1999 in order to overcome the foreign exchange crisis and activate the transportation industry and thereafter, the number of trucks rapidly increased compared to traffic leading to outcomes including deteriorated business conditions such as reduced load efficiency. The government has been making efforts to regulate supply and demand by way of switching freight transportation business reporting systems back to the licence system in January 2004 and freezing the number of trucks newly entering but the supply-demand imbalance of freight transportation has not yet been relieved and it is estimated that as of 2008, there was an oversupply of 15,000 trucks.

In the case of transportation businesses, ratios represented by transportation revenues and consignment revenues are almost 50% each and in the case of businesses that operate both transportation business and brokerage business, revenues are composed of transportation revenues 45.4%, brokerage revenues 22.4% and consignment revenues 30.8%. The fact that both dedicated transportation businesses and businesses that operate two types of business obtain more revenues through brokerage or consignments than through firsthand transportation indicates that the transaction structure in the market is not normal and that multi-level structures are prevailing. In addition, with regards to methods to secure transportation volumes, although both dedicated transportation businesses and businesses that operate two types of business receive around 75% of the volumes directly from shippers, they handle not all the volumes firsthand and thus it is assumed that freight is transported after going through many transaction steps.

The offline freight transportation market has complex transaction structures between shipper businesses, subsidiary logistics businesses, brokerage businesses, transportation businesses and car owners. In the domestic freight transportation market, dedicated consignment transport businesses represent the majority of participants accounting for around 60% of all transport businesses and general trucks transacting with these businesses still account for a large percentage of all general trucks although this percentage is in a trend of decreasing with around 70% in 2000 and around 60% in 2009.

The optimum transaction structure in the freight transportation market is a transaction structure where shippers having transportation volumes give freight to transportation businesses that are in charge of transportation and the transportation businesses deliver the freight to the final consignees. However, to review the current transaction structures in the freight transportation market, unnecessary transaction stages are added as transportation businesses and brokerage businesses receive volumes from initial shippers and they do not transport the volumes through their capacity but consign part or all of the volumes to other transportation businesses.

As a fundamental cause of the 2008 freight transportation refusal situation, many journalists pointed out truck transportation business transaction structures and although the pretext of the freight transportation refusal situation was increased transportation costs due to soaring oil prices, behind the scene, the transaction structures that was structurally irrational acted. As shown in Figure 2, the current freight transportation transaction structures in Korea are composed of at least four stages from the consigning of freight transportation to the delivery to the customer and 1-2 brokerage businesses are added between transportation businesses and truck owners. As taking 5~10% fees by stage has been established as practice, actual truck owners are receiving only 60~75% of the entire transportation charges.

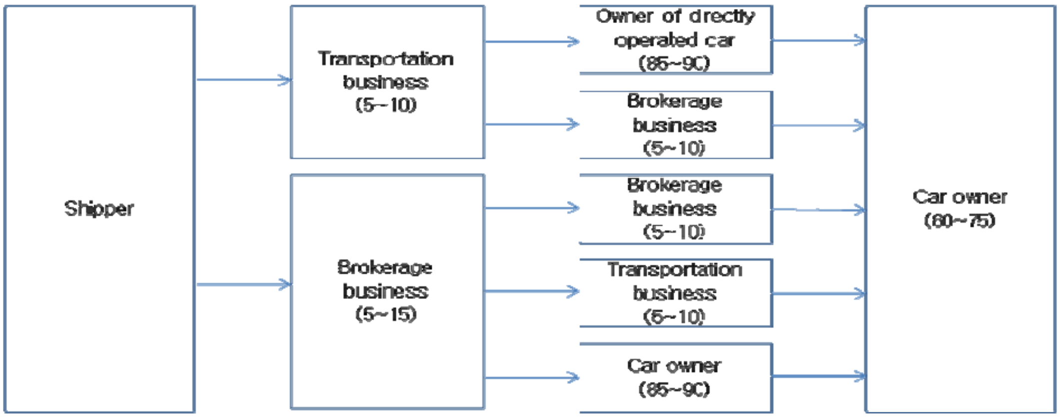


Figure 2. Freight Transportation Transaction Structure

2.3. Present State of The Online Freight Transportation Market and Related Problems

Recently, thanks to the development of Internet based on information technology and the activation of e-commerce, online based freight transportation services have appeared in the freight transportation market too and these services are perceived as one of the ways to fundamentally solve problems in truck transportation business such as multi-level transaction structures, high empty car ratios and supply-demand imbalance. Although the services at the beginning were in the form of simply posting freight or empty car information on notice boards, the services have been gradually developing into services with diverse forms and functions by subject being provided and along with the development of e-commerce, online freight transportation services are recognized as services that can immediately respond to freight transportation demand. In addition, it can be considered that online freight transportation transactions services are quite necessary in converting the unnecessary multi-level transactions which are thought to be the cause of the freight transportation refusal situation into direct transactions to solve related problems and reducing empty car ratios through sharing freight transportation information.

The online the freight transportation market can be considered as a sort of e-commerce markets and among them, it can be considered as an e-Marketplace where there are complex transaction relations of N:M participated by multiple suppliers and buyers. The characteristics of e-Market places are changing the complex transaction processes that have been followed offline into rational and efficient processes and epochally reducing transaction time and costs based on innovative processes and transparent transactions.

In the case of online based freight transportation services, car owners, shippers and transportation businesses gather together to transact while sharing freight transportation information and this online the freight transportation market can be considered as having the nature of markets where diverse entities gather together to transact not the linear nature of supply chains. The advantages of these online mediating services include, first, the reduction of transaction costs as costs for searching for cars and securing volumes are reduced through the sharing of large amounts of information, second, serving the role of preventing unfair transaction practice through direct connections between shippers and car owners made through the provision of transparency in volume transactions and third, shippers' trust in freight transportation that can be achieved by securing visibility of freight through the provision of real time car tracking functions. These online freight transportation transactions will be a good measure that can enhance the reliability of freight transportation, enhance the quality of transportation services, reorganize the multi-level transaction structures into efficient systems and improve the high empty car ratio issue.

As such, to fundamentally solve problems in truck transportation business such as multi-level transaction structures, high empty car ratios, supply-demand imbalance etc, introducing information technology into the freight transportation market and utilizing the technology can be a measure. However, to review the present state of online freight transportation transactions in Korea through the truck transportation . brokerage business survey report(2009) by the Korea Transport Institute(2009), only 3.1% of the businesses have a home page and thus it can be seen that their information system levels are quite lower compared to national information system levels and that as a result, the percentage of online transportation . brokerage transactions stayed at 3.4% and thus online transactions have not been activated. In addition, in the case of the capital or asset evaluated amounts of these businesses, 77.5% of them show less than 100 million won and 97.1% of them show less than 500 million won and thus it can be seen that most of businesses in the domestic freight transportation market are operated with low information systems and in petty sizes.

Table 2. Present State of Online Transportation/Brokerage Transactions

Division	Year	Online transportation transaction			Online brokerage transaction		
		Available	**Not available**	**Total**	**Available**	**Not available**	**Total**
All	2006	2.1	97.9	100.0	1.8	98.2	100.0
	2007	4.9(5.8)	95.1(94.2)	100.0	6.5	93.5	100.0
	2008	3.4(4.0)	96.5(96.0)	100.0	1.5	98.5	100.0

Truck transportation/brokerage business survey report, the Korea Transport Institute, 2009

3. LITERATURE REVIEW FOR E-MARKETPLACE SUCCESS ELEMENTS

This chapter tries to study success elements of e-Marketplace through review of previous researches through conceptual classification of the simply listed success elements. This study referred to existing domestic research literatures of Kim Bum-Yul(2000), Kim Sang-Soo(2001), Chun Myeong-Joong(2001), Seo Chang-Gyo(2001), Jung Dae-Yool(2002), Sung Haeng-Nam(2003), Lee Yoon-Ho(2002), Jung Seung-Ryul(2005), Hong Ji-Hwan(2006), and Philip Rosson & Charles Davis(2004), Warren D. Raisch(2001), Follit(2000), and Collins & Schefter(2004) in foreign countries. Success elements were classified into organizational classification, systematical classification, and political classification and the contents are follows;

Element Con		**tents**
OrganIzatIonal	Management's will	Management's mind, Management's interest and support, CEO's innovative minds, CEO's will, CEO's support
	Employees' participation	Positive participation of employees, The education level of employees
	Innovative minds	Innovation, Positive accommodation of new business trend, Customer-oriented mind, Provision of differentiated elements to participating companies in business, Immediate response to customers, Customer-oriented contents
		Improvement of business process, Creation of new company structure
		Creative and flexible managerial way of thinking
	Strategy, Vision	Strategic IT plan, Clear vision and plan for e-business
	Communication	Communication between companies, Management of communication, Establishment of virtual community, Utilization of various ways of communication, Exact communication, Communication with users, Sharing of information
SystematIc	Reliability	Business reliability, Transparency and honesty, Reliable service, Reliability of information
		Site security, Protection of personal information, Securing informational complementarity
		Ensuring exact and safe business
	Convenience, Efficiency	Increase of market efficiency, Efficiency of ordering, shipping and payment procedure, Site convenience, Business efficiency
	Connection, Integration	Connection with SCM, Immersion into business relationship, Mutual connection, Possibility of integrated support, Integration of supply chains, System integration
	Provision of information	Provision of various items, Provision of detailed information on product price, Provision of detailed information on product specification, Quality of information, Basic functions of the site
	Cost reduction, Income generation	Low price, Discovering the elements of continuous cost reduction, Suggesting model for generation, Maximization of participants' profit and balance of income, Creation of new types of customers, Securing new customers

	Supply capability	Product supply capability, Sufficient number of suppliers, Power of suppliers, Business initiative, Continuous data supply, Provision of information, Amount of information
Political	Governmental support	Governmental policy to positively support e-Marketplace, Governmental support, Laws and regulations related e-Marketplace, Strategic partnership
Others		Increase of purchasing power, Differentiation of business model, Excellent marketing ability, Suitability of operating body, Learning from successful company, Development of value-added service

On the basis of research on existing literatures, this study organized the success elements of e-Marketplace in following table, which shows that management's will, innovative mind, communication, reliability, connection/integration, cost reduction, and supply capability which have high frequency have positive effect on activation of e-Marketplace. Innovative mind is a concept which includes creative and flexible way of management as well as new company structure and improvement of processes, and reliability is a concept which includes not only systematic reliability but also safe and exact business between business bodies.

Researcher Research Item	Kim Bum-Yul	Maeil Business Newspaper	Kim Sang-Soo	Chun Myeong-Joong	Seo Chang-Gyo	Jung Dae-Yool	Sung Haeng-Nam	Lee Yoon-Ho	Jung Seung-Ryul	Hong Ji-Hwan	Bakos	B&W	S&W	Datacomm	Ramsdell	Wider	C&S	IDC	BCG	Follit	Harmanek et al	Warren D Raischh	Qizhi et al	P&C
Management's Will																								
Employees' Participation																								
Innovative Mind																								
Strategy, Vision																								
Communication																								
Reliability																								
Convenience, Efficiency																								
Connection, Integration																								
Provision of Information																								
Cost Reduction, Income Generation																								
Supply Capability																								
Governmental Support																								

4. INTERVIEW SURVEY

4.1. Interview Survey Method

In-depth interview was conducted on experts with many years of experience in freight forwarding business to examine the situation and problems of freight market and the ways of activating e-Marketplace. In this interview, questions were asked of consignors, car owners and experts with more than 15 years of experiences among those engaged in freight forwarders and shipping agents for their opinions on governmental policy related to freight forwarding market, the necessity and the ways for activation of e-Marketplace. In the first stage of the interview, on-line interview was conducted with the prepared questionnaires and, in the second stage, direct man-to-man interview was conducted to ask supplementary questions on the problems discovered in the first stage. Man-to-man interview was taped by a recorder and organized in documents by making script of each question and answer. The standard for selection of interviewees was limited to those in the position of managers or above working for consignors, car owners, freight forwarders, and shipping agents, and 1 researcher in relevant research institute was selected as research object to guarantee the perspective of the third party. The first on-line interview was conducted on experts from November to December 2009, and total 8 experts participated in the second man-to-man interview from January to February 2010. In order to guarantee anonymity, the participants in this study excluded specific personal information such as name and personal details except general characteristics.

4.2. The Result of Interview Survey

SK Netruck, as a freight forwarder and shipping agent, is providing the service of connecting consignors and car owners. On-line interview was conducted on the persons in charge of Pusan and Incheon branch offices of SK Netruck over the phone, where their opinions on on-line freight forwarding business could be heard. In the first place, as for the necessity of on-line freight forwarding, they had opinions that on-line business needs to be progressively expanded because it can reduce unnecessary procedures of business through direct connection of consignors with car owners, increasing the profit of car owners accordingly. As for the way of activation, they showed the opinion that promotional efforts through broadcasting or advertisement are needed because on-line freight forwarding is not yet familiar to the public, and that reliability guarantee system such as certification system for car owners will be necessary to reduce the worries of consignors when they request for the freight forwarding from car owners. As for the governmental policies, they said that the opinions of relevant parties should be continuously reflected in governmental policies rather than one-sided policies without reflecting the opinions of those currently working in freight forwarding business, and that there should be a main agent which plays the role of arbitrator between the policy makers and those currently working in the business.

And the result of interview with 'Hwamullmall', an on-line shipping agent, showed that they have opinion that freight forwarding-related problems can be solved through on-line business in the area of the necessity of on-line freight forwarding. They considered that it is essential in the first place to secure the participation of many users and sufficient volume of freight in the market in order to activate on-line freight forwarding business, and particularly that it is necessary to encourage on-line freight forwarding business under the initiative of government, and there was also an opinion that it would be a better idea to treat governmental supplies of PPS transparently through on-line freight forwarding. With regard to governmental policy of 'direct forwarding responsibility system' which allows the freight forwarding only for the companies with their own cars, the interviewees raised the necessity of reexamining the policy because the shipping agents, most of which have no cars of their own, can't conduct the freight forwarding.

Nextly, interview was conducted on car owners by visiting National Freight Car Owners Federation to hear the opinions of freight forwarding car owners. The car owners agreed on the necessity to solve the problems such as multi-level business structure in freight forwarding market, but also showed different opinions on the ways to solve the problem. First of all, acknowledging the necessity of on-line freight forwarding, they showed the opinion that the problem of multi-level can be solved and the profit of car owners can be enhanced by changing the freight forwarding structure through on-line business system. Those holding above opinions showed their intention to change to on-line system if practical profit is really coming to them, and that they are willing to use the terminals for on-line business even through education though they may have difficulty using the terminals. However, some of them were reluctant to change to on-line system while recognizing the problems of freight forwarding market because they are accustomed to off-line business system, which can be considered as their unwillingness to change to unstable and new market of on-line system because existing freight forwarding market is comparatively stable despite some of it's contradictory aspects.

Table 2. Result of Interview with Each Forwarding-Related Party

Question	Forwarding Company	Shipping Agent	Car Owners	Consignors
Necessity of e-Marketplace	•Business procedures can be reduced through e-Marketplace and the profit of car owners be increased accordingly	•On-line business can designate the freight in advance and can arrange freight forwarding with sufficient time	•Necessity of on-line business to solve the problems of multi-level and rate of empty vehicle	•Forwarding cost reduction can be achieved through securing stable forwarding company and market competition

Ways to activate e-Marketplace	•More participation of individual car owners can enhance the sense of belonging, giving reliability to consignors •Carrying out certification system for car owners to enhance reliability	•Participation of many users and sufficient exchange of freight information •Encouraging on-line business under the initiative of government •Securing the reliability of safe and timely delivery of freights, considering the situation of consignors •Formation of reasonable price through on-line	•Securing reasonable cost of forwarding •Stability of demand and supply of forwarding such as stable supply of vehicles and timely delivery	•Participation of reliable forwarders •Securing the stability (insurance, etc) of freight forwarding
Opinions on Governmental Policy	•The policy is carried out in desirable direction, with some insufficient reflection of opinions of those currently working in the business	•The policy of freight forwarding only by those who have their own cars means that the shipping agent can't conduct freight forwarding •Re-examination of the policy is necessary	•The intention of the policy is agreeable, but it doesn't properly reflect the current business situation •There is a difficulty in communication between the government and those currently working in the business, so opinion-arbitrating party is necessary	The policy is carried out in desirable direction, with some insufficient reflection of opinions of those currently working in the business

When rounding up the interview result, each party recognized the problems of freight forwarding market, and generally agreed on the necessity of on-line freight forwarding business as a way of solving problems. As the reasons for inactivated on-line freight forwarding business, the result showed the insufficient freight volume due to low rate of participation of users resulting from insufficient promotion of on-line system, the tendency to stick to and depending on existing off-line market where they can be continuously guaranteed of certain profit, and the low reliability on new on-line business market. Some of the interviewees even showed the tendency of settling for the present rather than showing the will for improvement even though they recognized the unreasonable situation of current market.

5. WAYS TO ACTIVATE E-MARKETPLACE OF FREIGHT FORWARDING MARKET

On the basis of general e-Marketplace success elements and the result of interview conducted on each party in freight forwarding market in previous studies, ways to activate e-Marketplace of freight forwarding market were organized as follows;

Division	Ways for Activation
Organizational	Arrangement of channel for smooth communication between the parties
	Management's support and consciousness reform for innovation
	Reinforcement of reliability by securing specialized freight forwarding companies
	Support for equipment introduction expenses for advancing into on-line market
	Positive promotional activities targeting the organizations and associations of each party, informing that there will be benefit arising from the participation in on-line market
	Reinforcement of freight supply capacity by securing the reliability from consignors
Systematic	Development and distribution of easy-to-use terminals
	Education on the users
	Provision of various additional services
Political	Arrangement and management of qualification standards for on-line market participants to guarantee the quality of services
	Participation of only those companies which satisfy the qualification standards to enhance the reliability

	Encouraging the participation through direct forwarding responsibility system and certification system for freight information network
	Provision of incentives to companies which conduct on-line forwarding business
	Increasing the demand by positively utilizing on-line business in procurement area

Activation elements were classified into organizational, systematic, and political aspects, and the analysis of each element showed the necessity of consciousness reform of members for arrangement of foundation for on-line business, and the necessity for smooth communication between the parties and for promotional activities to positively inform the advantages which can be obtained through on-line business. The analysis of systematic aspect showed the importance of development of terminals which can be easily used by the users in operating the new on-line system, and also the importance of education. And it also showed the necessity of encouraging participation of users by providing various services. In political aspect, positive support from government was found to be necessary, and particularly, it was found necessary to arrange and strictly manage the qualification standards for participants in order to create the frames to guarantee the quality of services provided and to create the environment of on-line freight forwarding business market. In addition, if on-line business can be used in governmental procurement business which occupies considerable volume, certain level of demand can be constantly secured, achieving the additional result of increasing on-line business demand.

6. CONCLUSION

This study conducted empirical analysis to examine which elements should be considered first in order to activate e-Marketplace targeting forwarding companies, shipping agents, and car owners in freight forwarding market. It was found in this study that communication, innovative mind, reliability, cost reduction, and profit generation are essential for activation of e-Marketplace, through the process of conceptually classifying simply listed success elements by suggesting and analyzing the success elements through the examination of previous studies. On the basis of previous studies, this study divided the elements for activation into management's will, innovation, communication, reliability, connection and integration, cost reduction, profit generation, and supply capability, and on the basis of interview results conducted on each party of freight forwarding market, it was confirmed that profit maximization, convenient use, securing freight volume, and reliable business are important elements for activation.

It is considered that this study will serve as a foundation for development of future on-line forwarding business, suggesting important considerations in introduction and establishment of e-Marketplace of freight forwarding market which is expected to continuously expand in the future. But this study of e-Marketplace of freight forwarding market is just in the beginning. Accordingly, this study is focused on situation survey through case study and interview, so continuous study on analysis of actual business case of e-Marketplace and empirical verification of freight forwarding market should be conducted in the future.

ACKNOWLEDGEMENTS

This work was supported by National Research Foundation of Korea- Grant funded by the Korean Government (KRF-2009-32A-B00090)

7. REFERENCES

Bakos, Y.(1991). Information Links and Electronic Marketplaces : The Role of Interorganizational Information Systems Vertical Markets, *JMIS*, Vol. 8, No. 2, pp. 31 ~ 52.

Benjamin, R. and Rolf Wigand(1995). Electronic Markets and Virtual Value Chains on the Information Superhighway, *Sloan Management Review*

Churl-Min Kim(2004). *Empirical Research and Revitalization Strategies of Internet Based Cargo Transportation Business*, YoungSan Univ. Research Report

Collins M., and P. Schefter(2000). The Narrow Path to Victory for B2B Exchanges, *Harvard Management Update*, Vol. 5, No. 9,

Datacomm Research Company(1999). *Portals to Profit : E-Commerce Business Models and Enabling Technologies*

Follit, E.(2000). *The Keys To E-Transformation*, Informationweek.com

Jung-A Kim, Ha-Jin Hwang(2006). An Empirical Study to Identify Success Factors of B2B e-marketplaces, *Journal of Internet e-Commerce Research*, Vol 6 No.2

Ministry of Land, Transport and Maritime Affairs. http://logistics.mltm.go.kr

Yong-Jin Kim, Seung-Bum Ahn, Jung Ung(2005). Strategies to Stimulate On-line Marketplaces for Truck Freight Exchange Services, *Journal of Logistics Research*, Vol.13 No1

Philip Rosson and Charles Davis(2004). Electronic Marketplaces and Innovation : the Canadian Experience, *International Journal of Information Technology Management*, Vol. 10, No. 10, pp. 1 ~ 18.

Qizhi Dai and Robert Kauffman J.(2002). Business Models for Internet-Based B2B Electronic Markets, *International Journal of Electronic Commerce*, Vol. 6, No.4, pp. 41 ~ 72.

Sang-Soo Kim, Jong-Hae Ha(2001). An Empirical Study on the Critical Success Factors of MRO e-marketplace, *Journal of IT and Database*, Vol 8 No2

Sculley A. B. and W. W. A. Woods.(1999). *B2B Exchanges : The Killer Application in the Business-to-Business Internet Revolution*, ISI Publications

The Korea Transport Institute(2009). *Truck transportation/brokerage business survey report*

Wider C.(2000). *E-Business : What's the Model?*, Planet IT

Session D4: Maritime 2

·Day1: Sep. 15 (Wed.)

·Time: 16:40 - 18:00

·Chair: Namkyu Park

·Room: Grand Ballroom, 5F

LOGMS

The 1st International Conference on Logistics and Maritime Systems

September 15th–17th, 2010 Busan, Korea

Lin

THE LOGISTICS OF FACILITY LOCATION ON SPHERICAL SURFACES

Jovin J. Mwemezi

Department of Logistics Management and Engineering
Shanghai Maritime University
1550 Pudong Avenue
Shanghai, China
e-mail: mwemezijm@yahoo.com

Abstract: Fundamental to modelling of location decision is some measure of proximity. In measuring proximity, point-to-point d istances are often u sed. Distance m easurements used i n facility location are pred ominantly b ased on Rectilinear and Eu clidean distances. Th is paper seek s to pr esent an al ternative distance m easurement “great circle distance” which re presents t he shortest pa th on spherical s urface. Great circle distances takes int o conside ration the geometrical reality o f the spherical Earth and will present an option to wid ely held notion that travel over water can be exactly modelled by Eucl idean di stances. The need for geom etrical presentation of the sp herical e arth becomes very relevant when we tak e i nto consideration an ev er i ncreasing facility locatio n at sea where great ci rcle trav el can be practised. Facilities being located at sea include oil rigs, mobile drilling units and dynamically positioned units. The use of “Great ci rcle di stances” o pens u p a nother ave nue for convergence of Na vigation a nd S pherical T rigonometry i nto advancement of l ogistics and facility lo cation. In th is p aper an ev aluation of sing le facility lo cation usin g great circle distances is made in order to demonstrate the concept.

1. INTRODUCTION

It is recognised that th e location of a facility d etermines an d has great influ ence on the distribution sy stem p arameters including time, costs and e fficiency of t he system (Sule, 2001) . As su ch, optimal location of th e facility is essen tial to attaining im poroved flow of g oods an d serv ices to cu sstomers serv ed by th e facility. In cho osing t he lo cation of a facility b oth qualitative and quantitative facto rs are tak en in to account i ncluding av ailability o f land , prox imity to raw materials o r mark et, av ailability o f u tilities an d tran sport facilities as well as so cial, eco nomic an d political factors (Zarimbal, 2009; Melo *et al*, 2005). Distance or proximity is on e of th e important metric wh ich many d ecision makers seeks to optmise through m inimization of the mean (or t otal) distance as in the median concept or minimization of the maximum distance as in the centre concept (Schilling *et al,* 1993; Rodríguez-Chía, 2010).

Though distance is a wel l known parameter, its determination in certain settings could be challenging like finding out distance between positions de fined by l atitudes an d l ongitudes on Earth. C onsidering, t he s pherical nat ure of t he Earth it is e vident that distance modeling in facility location that takes into account this fact will be an improvement on the current practice dominated by Euclidean and Rectilinear models which are best suited to planar surfaces. This paper seeks to present an alternative distance measurement based on “great circle distance” which represents the shortest path on spherical surface. The need for geometrical modeling of distance of the spherical earth becomes very relevant when we take into consideration an ever increasing facility location at sea for h arnessing natural resources including oil rigs, mobile d rilling units and dyn amically p ositioned units. Un like travel on land wh ere p hysical b arriers h ave to be avoided, it is practical to travel along the great circle path during open sea navigation.

Logistics has borrowed theories from many other disciplines of study like marketing, mathematics and psychology (Stock, 1997; Sach an and Datta, 2005; Gamme lgaard, 2 004). The use of “Great circl e di stances” opens up another avenue for borrowing from navigation and spherical trigonometry into advancement of logistics and facility location. In this p aper single facility lo cation based on great circle distances is ev aluated in t he p rocess of de monstrating and applying the concept.

2. DISTANCE FUNCTIONS IN LOCATION PROBLEM

Zarimbal (20 09) clearly affirm that th at th e d istance functions p lay an im portant ro le in facility lo cation prob lems. He identifies di fferent distance functions used in location problem with each ha ving its own domain, adva ntages, a nd disadvantages. He defines distance as a numerical description of how far apart objects are at any gi ven moment in time and m ay refer t o a physical l ength o r a period of t ime. While m aking l ocation d ecisions, network desi gn a nd optimization; the distribution of travel distances among the service recipients (clients) is an important issue (Fernandez et al, 2002).

Based on the work Zarinbal (2009) we note that Euclidean and Rectilinear distance accounts for more than 63 percent of distance functions used in location problems. Euclidean distance assumes that one can travel almost directly from one station to another following a straight line as shown in figure 1 (Montreuil, 2008; Melachrinoudis and Xanthopulos, 2003).

Rectilinear distances are applicable when travel is allowed only on two perpendicular directions such as North–South and East–West arteries as shown by the dotted line in figure 2. This distance is also popular among researchers because the analysis is usually simpler than employing other metrics (Drezner and Wesolowsky, 2001). The rectilinear distance is also called Manhattan or Taxicab Norm distances; because it is the distance a car would drive in a city lay-out in square blocks. Apart from these two dominant distance functions, other distance used in location problem includes aisle distance, distance matrix, minimum lengths path, Hilbert Curve, Mahalanobis distance, Hamming Distance and Chebyshev Distances (Klamroth, 2002; ReVelle and Eiselt, 2005). Klamroth (2002) groups these distances into multi-parameter round norms, block norms and polyhedral distances.

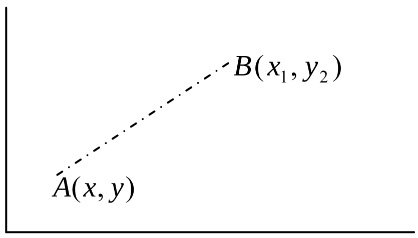


Figure 1: Euclidean Distance

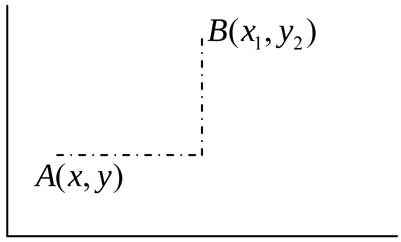


Figure 2: Rectilinear Distance

Thus Euclidian distance between two points A and B with co ordinates A (x, y) and B (x_i, y_i) can be expressed mathematically as;

$$d(s) = \left[(x - x_i)^2 + (y - y_i)^2\right]^{\frac{1}{2}} \tag{1}$$

Rectilinear distance $d(s)$ between A (x, y) and B (x_i, y_i) co-ordinates can be expressed as:

$$d(s) = |x_i - x| + |y_i - y| \tag{2}$$

2.1 Deficiency of Current Distance Modeling in Location Problem

In a realistic environment the choice of a suitable distance function plays a crucial role for a good estimation of travel distances (Klamroth, 2002 and Sminchi-Levi, 1997). In reality we are located on spherical Earth with our addresses defined by the intersections of latitudes and longitudes. Precise geographic locations can be achieved by using a geographic information system (GIS) and other satellite based systems like the Global Positioning Systems (GPS) and Glonass as well as navigation charts (Manley, 2008). Grid systems can also be used to model location and travel distance but suffer from having limited use as most of them are established based on national grid reference system hence inappropriate to evaluate facility location and networks that spans across the borders of countries with different grid reference system (Wikipedia 1).

Bramel and Sminchi-Levi (1997), Klamroth (2002), Drezner and Wesolowsky (2001), and Zarinbal (2009) assert that air travel and travel over water can be exactly modeled by Euclidean distance. However, this suggestion disregard the fact that air travel and sailing at sea is made over a spherical surface whereas Euclidean modeling simply measures the distance that would be obtained if the distance between two points were measured with a ruler (Zarinbal, 2009). Using the Pythagorean Theorem (as in Euclidian Distance) and Spherical Trigonometry principles reveals disagreement between the measurements and the calculations of the sides and angles. In fact, the sum of the angles in spherical triangle is greater than the 180 degrees which is always measured in planar triangles (Ross, 2002; Wikipedia2). The discrepancy between the distances measured based on Euclidean and those based on spherical trigonometry becomes greater, the further apart the locations are from each other (Ross, 2002). Modeling distance of air travel or ocean navigation using Euclidean distances is in principle asserting that such travel is made through the interior of the sphere which is not the case. This anomaly can be corrected by use of spherical trigonometry as proposed in this paper.

2.2 Great Circles in Distance Modeling

Based on the work of Ross (2002), Frost (1988) and Earl et al (1999) we note that Trigonometry and spherical trigonometry were primarily developed for and used in astronomy, geography, and navigation. Spherical trigonometry was developed to describe and understand applications involving triangles on spheres and spherical surfaces. Spherical trigonometry offers a realistic representation of the Earth surface which is spherical in nature and is widely used in other discipline of studies but its potential particularly the use of great circle distances remains untapped in logistics. The potential areas for application of spherical trigonometry concepts include but not limited to hub-and-spoke network design and facility location at sea like oil and gas rigs.

Measurement of distances in spherical trigonometry is based on solving spherical triangles whose sides form arcs of great circles (Das *et al*, 2001). As in figure 1, great-circle arcs form the sides of a spherical triangle, and where two arcs intersect, a spherical angle is formed. In other words, the arc lengths are a measure of the angle they subtend at the center of the sphere, and the spherical angles between the arcs are a measure of the angle at which the planes that form the arcs intersect. On the Earth, the equator and circles of longitude are natural great circles. Likewise, any circular path around the Earth that cuts it into two equal hemispheres is a great circle. Spherical trigonometry involves relationships between the arc lengths (sides) and the spherical angles between the arcs.

Studies have shown that the shortest distance between any two positions on the earth's surface lies along the arc of the great circle joining these two positions. Thus on a spherical surface, a great circle path, often called a geodesic, is always the shortest path between two points (Ross, 2002). And between any two points on a sphere which are not directly opposite to each other there is a unique great circle (Wikipedia).

In recognition of the fundamental difference between spherical geometry and Euclidean Geometry it is apparent that the equations for distance take different forms in these two domains of knowledge. Fundamentally, the distance between two points in Euclidean space is the length of a straight line from one point to the other while in spherical geometry straight lines are replaced with geodesics or great circle paths.

While positions of the geographical places can relatively be easily determined based on existing maps or global positioning systems like GPS and Glonass the calculation of the great circle distance and thus the shortest distance between places needs a formula. By using a system of co-ordinates of longitude and latitudes the distance along the great circle can be determined by solving the quantities of the resultant spherical triangle formed by the intersection of three great circles (Frost, 1988) namely:

a) The great circle arc joining the two positions (arc c in figure 1)
b) The meridian (longitude) through position 1 (meridian joining C and A in figure 1)
c) The Meridian (longitude) through position 2 (meridian joining C and B in figure 1)

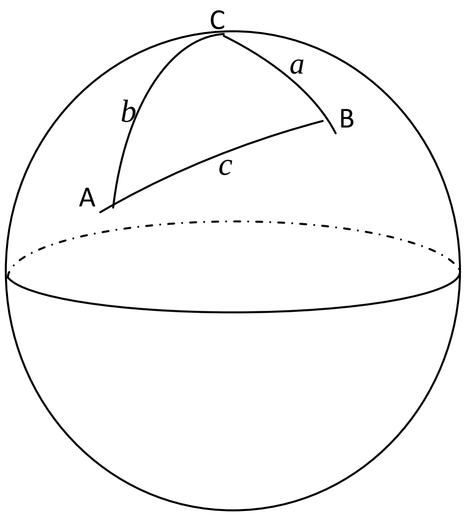


Figure 3. Spherical Triangle

Such spherical triangles and shortest distance between geographical points are solved by using the haversine formula (Bell *et al*, 2010) as shown in (3), (4) and 5. Thus in spherical triangle ABC in figure 1 above, given CA or b, CB or a and angle C, the haversine formula to solve arc length AB or c is expressed as:

$$hav(dist) = hav(dlong)\cos lat(A)\cos(latB) + hav(dlat) \tag{3}$$

Or

$$havAB = hav(C)\sin(a)\sin(b) + hav(a-b) \tag{4}$$

Where:
$hav.dist$ – Haversine of distance between position A and B
$hav.dlong$ – Haversine of the difference between the longitudes through position A and B respectively
$\cos lat.A$ – Cosine of the latitude through position A
$\cos lat.B$ – Cosine of the latitude through position B
$hav.dlat$ – Haversine of the difference between latitudes through position A and B respectively

Alternatively, great circle distance can be calculated by finding the interior spherical angle between the two points and then multiplying that angle by the radius of the earth. Thus the length of the side of the spherical triangle (distance S) in figure 4 is given by:

$$S = \alpha r \tag{5}$$

$S =$ Arc length (great circle distance on the sphere)
$r =$ Radius, in this case the radius of the earth which is 6,371.009 km or 3,958.761 miles or 3,440.069 nautical miles
$\alpha =$ Central angle measure

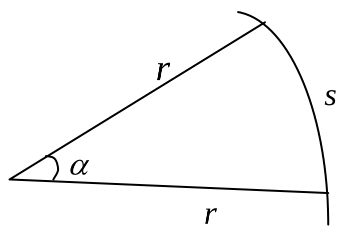


Figure 4. Arc length (S) and Central angle

Based on the haversine formula the central angle in radians is expressed as

$$\alpha = 2\arcsin\left(\sqrt{\sin^2\frac{\Delta\theta}{2} + \cos\theta_1\cos\theta_2\sin^2\frac{\Delta\varphi}{2}}\right) \tag{6}$$

Where: $\alpha = Interior \quad Spherical \quad Angle$
$\Delta\theta = \theta_1 - \theta_2 = difference \quad in \quad latitude.(dlat)$
$\theta_1 = Latitude \quad at \quad position 1$
$\theta_2 = Latitude \quad at \quad Position 2$
$\Delta\phi = Difference \quad in \quad Longitude.(dlong)$

In general the different forms of the haversine formula can be deduced from the law of cosine for spherical triangle. For spherical triangle ABC in figure 1 the cosine rule is stated as:

$$\cos(c) = \cos(a)\cos(b) + \sin(a)\sin(b)\cos(C)$$
$$\cos(a) = \cos(b)\cos(c) + \sin(b)\sin(c)\cos(A)$$
$$\cos(b) = \cos(a)\cos(c) + \sin(a)\sin(c)\cos(B)$$

In order to deduce the distance from the haversine formula the haversine tables are used. Value in (4) can be computed directly using calculators or some software and programs which have been developed that solves the distances directly utilizing the excel capabilities.

The haversine formula we will assume that the Earth is a perfect sphere, even though it really isn't but somewhat ellipsoidal at the poles. To correct this anomaly a more complicated formula known as Vincenty's formula (equation 7) was developed (wikipedia; Jenness, 2008). Except for the antipodal point (points on the sphere directly opposite to each other), the haversine formula gives accurate distance. For demonstration in this paper the haversine formula has been used.

$$\alpha = \arctan\left(\frac{\sqrt{(\cos\theta_2 \sin\Delta\varphi)^2 + (\cos\theta_1 \sin\theta_2 - \sin\theta_1 \cos\theta_2 \cos\Delta\varphi)^2}}{\sin\theta_1 \sin\theta_2 + \cos\theta_1 \cos\theta_2 \cos\Delta\varphi}\right) \quad (7)$$

3. APPLICATION OF SPHERICAL TRIGONOMETRY IN LOCATION PROBLEM

The starting point in determining the optimum location is to find the centroid/centre of gravity of the spherical polygon under con sideration (Jenn es, 2008) as t he in itial co ordinate o f th e new facility. Calcu lating a cen troid fo r sph erical surface is complex and still being st udied, however, it is sim ilar in concept as calculating of planner surface. The main difference i s that l ongitudes and l atitudes a re not s o much of c oordinate b ut rat her di rections from t he ce ntre of t he sphere. Si nce l ongitude a nd l atitudes ca n n ot be si mply adde d an d di vided as t he C artesian c oordinates can (Je nnes, 2008), we first convert them into radians for calculating the centre of gravity. In order to facilitate calculations by excel, the positions given in degree, minutes and seconds a re c onverted int o decimal places and ra dians (Pears on, 2009). Likewise, in order to take in to account the hemisphere in which the position lies we introduce negative values for South Latitudes and West longitudes.

3.1 The Algorithm for Apply Great Circle Distances in Facility Location

i. Express latitudes and longitudes given in degrees, minutes and seconds as decimal values
ii. Express West Longitudes and South Latitudes as negative values otherwise positive
iii. Express location coordinates as radians by first converting degrees, minutes and second into decimal then apply

$$Radians = \frac{\deg}{180} * \pi \quad (8)$$

iv. Determine the initial co-ordinate of the new facility (x^*, y^*) defined by centre of gravity formula

$$x^* = \sum w_i x_i \Big/ \sum w_i \text{ and } y^* = \sum w_i y_i \Big/ \sum w_i \quad (9)$$

v. Calculate the total distance (cost) from (x_i, y_i) to the optimal location using the haversine formula
vi. Make an iterative search of minimal total distance/cost based on initial position until no improvement is found

3.2 Application of the Concept Demonstrated

ABC Company Ltd has 8 oil and gas rigs located in the Gulf of Guinea –West Africa with coordinates as shown in table 1 (locations for demonstration purpose only). On reviewing its policy on distribution of supplies to the rigs, ABC plans to se nd s upplies once i n m onth t o a cent ral wa rehouse b y usi ng a ship. F rom t he ce ntral warehouse the s upplies ar e distributed to other rigs by smaller boats making 2 trips to rig no.1 and 6 and only one trip to all other rigs per week. The task is t o d etermine th e lo cation of t he cen tral warehouse that m inimizes the t ravel distance and therefore the distribution costs.

Table 1: Location of Oil and Gas Rigs

RIG	LOCATION (DEGREE)				LOCATION (DECIMAL)	
	LATITUDE		LONGITUDE		LATITUDE	LONGITUDE
1	01º 36' 15"	N	07º 37' 17"	E	1.6041667	7.6213889
2	02º 00' 27" S	0	6º 25' 50"	E	-2.0075	6.4305556
3	00º 40' 38"	N	09º 05' 26"	E	0.6772222	9.0905556
4	01º 04' 23" S	0	5º 19' 45"	E	-1.0730556	5.3291667
5	00º 03' 06" S	0	8º 07' 26"	E	-0.0516667	8.1238889
6	02º 53' 52"	N	08º 59' 18"	E	2.8977778	8.9883333
7	00º 28' 46"	N	06º 57' 30"	E	0.4794444	6.9583333
8	01º 05' 36" S	0	8º 27' 14"	E	-1.0933333	8.4538889

Solution:
We have noted that Great circle distance provides the shortest distance between two positions on the surface of the earth; hence great circle distances are used in solving this problem using the algorithm stated above.

The objective function: Minimize

$$S = \sum_{i=1}^{n} w\alpha_i r \text{ -----from (5) where } w = Trips$$

Step 1: Express location co-ordinates in decimal format (see Table1)
Step 2: Express West Longitudes and south latitudes as negative values otherwise positive (see Table 1 and figure 3)
Step 3: Express Location coordinates as radians (see Table 2)
Step 4: Determine the initial latitude (x^*) and initial longitude (y^*) as shown in table 2

Table 2: Calculation for Initial Location

Rig	Trips (w_i)	Latitude in Radians x_i	Longitude in Radians y_i	$w_i x_i$	$w_i y_i$	$x^* = \sum w_i x_i / \sum w_i$
1 2		0.027998	0.1330183	0.055996	0.2660367	= 0.0103585 radians
2 1		-0.0350375	0.1122344	-0.0350375	0.1122344	= 00º 35' 37"
3 1		0.0118198	0.1586601	0.0118198	0.1586601	
4 1		-0.0187284	0.0930115	-0.0187284	0.0930115	
5 1		-0.0009018	0.1417886	-0.0009018	0.1417886	$y^* = \sum w_i y_i / \sum w_i$
6 2		0.0505758	0.156876	0.1011515	0.313752	= 0.1354477 Radians
7 1		0.0083679	0.1214458	0.0083679	0.1214458	= 07º 45' 38" E
8 1		-0.0190823	0.1475482	-0.0190823	0.1475482	
	10			**0.1035853**	**1.3544773**	

Based on the calculation performed in Table 2, the initial optimal location of the central warehouse will be at 00º 35' 37"N, 07º 45' 38"E.

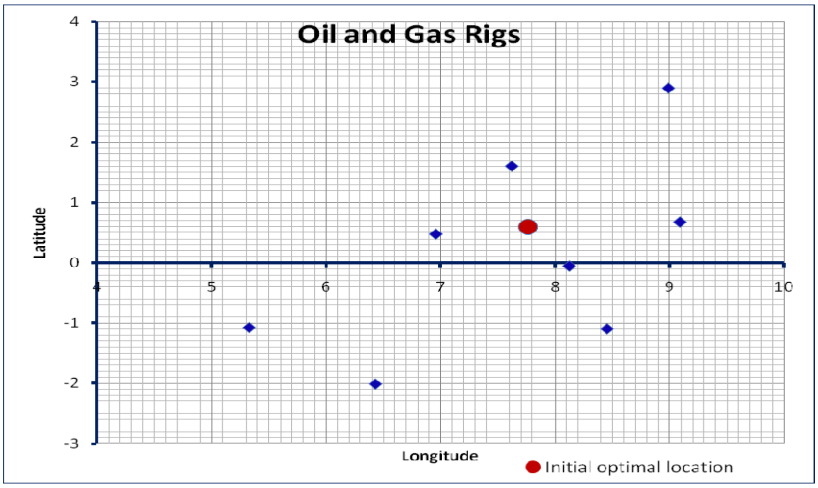


Figure 3: Location of oil rigs

Step 5: Calculate the total distance (cost) from (x_i, y_i) to the optimal location using the haversine formula:

Table 3: Calculation of Great Circle Distances based on Geographical Coordinates

RIG LO	CATION	(DECIMAL)	Trip	$S = \alpha_i r$	$w_i \alpha_i r$ (Km)
	LATITUDE	LONGITUDE			
1 1	.6041667	7.6213889	2	113.443	226.886
2 -	2.0075	6.4305556	1	324.831 32	4.831
3 0	.6772222	9.0905556	1	148.138	148.138
4 -	1.0730556	5.3291667	1	327.769	327.769
5 -	0.0516667	8.1238889	1	82.33	82.33
6 2	.8977778	8.9883333	2	290.295	580.59
7 0	.4794444	6.9583333	1	90.101	90.101
8 -	1.0933333	8.4538889	1	202.794 20	2.794
	x^*=0.5935	y^*=7.760583		$S = \sum_{i=1}^{8} w \alpha_i r$	1983.439

The total initial distance between the initial optimal location and all other stations is 1983.439 Kilometers. The initial position provides valuable in put for su bsequent iteration s an d d ecision m aking process tak ing i nto accou nt both qualitative an d qu antitative an alysis. If a new facility is to be con structed t he cen tre of grav ity (x^*, y^*) is an ideal location. Alternatively one of the existing facilities (rig) can be used for central warehouse. By inspection we note that rig 5 is closest to the centre of gravity, hence the ideal candidate for second iteration.

By lo cating t he wareh ouse at rig 5 we note th at th e to tal distribution di stance (c ost) becom es 2161.965 Km an increase of 178 .526 Km as compared to l ocating a new f acility at the cen tre of gravity. Th e additional annu al distribution cost related to the extra 178.526 Km per week needs to be compared to the annual fixed cost of establishing a new facility. This will help in establishing the tradeoff between establishing a new facility and locating a warehouse at the existing facilities. Iterations can be made for all the remaining facilities and evaluations made accordingly.

4. CONCLUSION

Analysis made in this paper shows that distances for facility location can be modeled more realistically by applying the great circle distances concept that take s into account the spherical nature of t he Earth we live in. T he contribution of this paper has been th e in troduction of an alternative appr oach to distance modeling for travel over water in place of Euclidean distance by developing an algorithm for deducing distances from geographical address defined by the grid of latitudes and longitudes and applying spherical trigonometry principles in the logistics of facility lo cation. As such, the application o f "Great ci rcle di stances" which i s very much used i n navi gation a nd Spherical Trigonometry will contribute to advancement of logistics and facility location by broadening the scope of the set of knowledge from which the logistics discipline borrows.

ACKNOWLEDGEMENT

This w ork was su pported i n part by Sha nghai Sci ence Commission P roject (No. 0 9DZ2250400), Shanghai Science Commission International Cooperation Project (No.09530708200), and Shanghai Municipal Science Commission Local University Cap ability Pro ject (No . 08170511300), Sh anghai Education Co mmission Leading A cademic D iscipline project (No. J50604) and Shanghai Science Commission Project (No. 10692103500).

5. REFERENCES

Bell, J.E. *et al.* (2010). Location optimization of strategic alert sites for homeland defense. *Omega*. Article in press

Bramel, J and Sim chi-Levi, D . (1997). *The Logic of Logistics: Theory, Algorithms, and application for Logistics Management.* Springer-Verlag, New York, USA

Das, P. et al. (2001). Spherical Minimax Location Problem. *Computational Optimization and Applications, 18,* 311–326

Earl, G.E et al (1999). *MUNRO'S Mathematics for Deck Officers.* . Brown, son & Furguson LTD, Nautical Publishers

Frost, A. (1988). *The Principles and Practice of Navigation.* Brown, son & Furguson LTD, Glasgow, UK

Gammelgaard, B. (2 004). Schools in log istics research ? A methodological fram ework for an alysis of th e discipline. *International Journal of Physical Distribution & Logistics Management, 34,* 479-491

Jenness, J. (2009). *Geodesic Tools: Extension for ArcGIS*. Jenness Enterprises

Klamroth, K. (2002). *Single facility location problems with barriers.* Springer-Verlag, New York, USA

Manley, P. (2008). *Practical Navigation for the Modern Boat Owner.* John Wiley & Sons Ltd, West Sussex, England

Melachrinoudis, E. and Xanthopulos, Z. (2003). Semi-obnoxious single facility location in Euclidean space. *Computers & Operations Research, 30*, 2191–2209

Melo, M. T. et al (2005). Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning. *Computers & Operations Research, 33,* 181–208

Montreuil, B. (2008). "Facilities Lo cation and Lay out Design", in Tay lor, G.D., *Logistics Engineering Handbook,* Taylor & Francis Group, New York, USA, 9.1 - 9.56

Pearson, C. H. (2009): Latitude, Longitude and Great Circles. http://www.cpearson.com/excel/LatLong.aspx

ReVelle, C. S. (2005). Location analysis: A synthesis and survey. *European Journal of Operational Research*, 165, 1–19

Rodríguez-Chía, A. M. *et al* (2010) Continuous Optimization: On solving the planar k-centrum problem with Euclidean distances. *European Journal of Operational Research*, Article in Press

Ross, D. A., (2002). *Master Math: Trigonometry*. Career Press, NJ, USA

Sachan, A . and Datta, S. (2005). Rev iew of su pply ch ain m anagement and l ogistics r esearch. International J ournal o f Physical Distribution & Logistics Management, *35*, 664-705

Schilling, D. A. *et al*. (1993). A Review of Covering Problem in Facility Location. *Location Science*, 1: 25-55

Stock, R. (1997). Applying theories from other disciplines to logistics. *International Journal of Physical Distribution & Logistics Management, 27,* 515-539.

Sule, D. (2001). *Logistics of Facility Location and Allocation.* Dekker Inc, New York, USA

Web GIS: Geodesic Tools**.** http://webgis.wr.usgs.gov/pigwad/docs/Geodesic_Tools_Manual.pdf

Wikipedia: Haversine formula. http://en.wikipedia.org/wiki/Haversine_formula

Wikipedia: Great circle distance. http://en.wikipedia.org/wiki/Great-circle_distance

Wikipedia: Grid reference system. http://en.wikipedia.org/wiki/British_national_grid_reference_system

Wikipedia: Spherical trigonometry http://en.wikipedia.org/wiki/Spherical_trigonometry

Zarinbal, M. (2 009). Distance Functions in Lo cation Problems, in Far ahani, R.Z and Masoud H ekmatfar, M. *Facility Location: Concepts, Models, Algorithms and Case Studies,* Springer-Verlag, Berlin Heidelberg, 5-18

SENSITIVITY MODEL OF TRANSHIPMENT CARGO BY SYSTEM DYNAMICS

Namkyu Park[1*], Chaekwan Lim[2], and Branislav Dragović[3]

[1]School of Port and Logistics Management
Tongmyong University
535, Yongdang-dong, Nam-gu,
Busan, Korea
e-mail: nkpark@tu.ac.kr

[2]Department of Distribution Management
Tongmyong University
535, Yongdang-dong, Nam-gu,
Busan, Korea
e-mail: cklim@tu.ac.kr

Maritime Faculty,
University of Montenegro, Maritime Transport & Traffic Division
Dobrota 36,
85330 Kotor, Montenegro,
branod@cg.ac.yu

Abstract: The aim of this paper is to identify impacting factors that have been affecting the increase of transshipment cargoes of port of Busan and to identify forecasted result using system dynamics. To clarity the reason why T/S cargoes have increased in the port of Busan, several steps are made as follows: The first step is to make a quantitative model for explaining the development of T/S cargoes during the last decade. To define dependent and independent variables for multiple regressions after testing variable significance is the second step. For this, data collection and the accuracy of validation have been done by the direct interview with the experienced officials in shipping companies of both domestic and foreign country. After validating the model with collected data, the final step is to find variables which are explaining the model the most. In conclusion, 2 variables were clearly identified as core factors that explain well the development of T/S cargoes in the port of Busan: 'Mohring effect' and total cost. It is strongly recommended, by an empirical study, that an incentive scheme be changed to a way which more feeder vessels rather than mother vessels can reduce their direct costs to call in the port of Busan. Based on regression analysis, sensitivity model for transhipment cargo is useful for dynamic forecasting in changing cost factor and mohring factor with time series.

Key words: Transshipment Port, Choice of Transshipment Port, Incentive, Busan Port Authority, SD(System Dynamics)

1. INTRODUCTION

Being hub port in the Northeast Asia area by pulling transshipment cargo is the vital issue of competing ports in order to revitalize their economy and to overcome the problem of under utilization of container terminal facility. The port of Busan, Shanghai, Ningbo and Hong Kong, Kaosiung and Yokohama in Northeast Asia ports has been the competitive relation in terms of transshipment attraction. This paper is to make port choice model in order to indentify for finding out impacting factors to transshipment cargo and sensitivity consequence from the model. In order to gain the research objective, several steps are designed. First step is to select quantitative model for explaining real phenomena about transshipment cargo share. Second step is to define dependent and independent variables for multiple regressions after the test of variable significance. On this step, data collection and the accuracy validation has been done by direct interview with the experienced official in shipping company in domestic and foreign country. As Northeast cargo markets have separate structures, the transshipment cargo from China, Japan or Southeast countries, individual market is to be considered independently. Third step is to validate the model using collected data, in order to reveal which variables explain the model in a good fit. Lastly, port choice sensitivity model is to be developed to find estimation according to changing independent variables.

The scope to find impacting factors is restricted within North China ports including port of Busan and port of Sanghai port in consideration of clarify objective and descriptive factors.

2. LITERATURE REVIEW

Hiroshi Ohashi (2005) studied the choice problem of air cargo transshipment airport in Northeast Asia. Based on a unique data set of 760 air cargo transshipment routings to/from the Northeast Asian region in 2000, this paper applies an aggregate form of a multinomial logit model to identify the critical factors influencing air cargo transshipment route choice decisions. The analysis focuses on the trade-off between monetary cost and time cost while considering other variables relevant for choice of transshipment airport. The estimation method considers the presence of unobserved attributes, and corrects for resulting endogeneity via a two-stage least-squares estimation using instrumental variables. The empirical results show that choice of the air cargo transshipment hub is more sensitive to time cost than the monetary costs such as landing fees and line-haul price.

Veldman and Bückmann (2003) analyzed earlier with respect to container port competition in Northwest Europe. They estimated demand functions for both the continental and the overseas hinterland of the West European major container ports and assessed the demand function for a port expansion project for the port of Rotterdam.

Veldman et al. (2005) estimated demand functions for a project to improve the accessibility of the Port of Antwerp by deepening the Scheldt River and thereby reducing waiting times for the tide and the ability to accommodate bigger ships. In both publications the parameters of a Logit Model were estimated with regression analysis and the demand function could be derived by systematically changing cost and assessing the resulting market shares.

Veldman et. al. (2008) studied to search significant factors for understanding the competitive position of transshipments ports and port choice elastics in the market of the Mediterranean. Statistical tests are applied using a 10-year time series of aggregate transshipment flows between 15 transshipment ports and 9 feeder regions. Tests of Logit Models with regression analysis show that variables such as feeder costs, mainline port access costs and Mohring effects are statistically significant.

Also Lirn et al. (2003, 2004) and Ng (2006) have analyzed the decision factors for transshipment port and have revealed that the cost of a shipping company, route accessibility, and time are important decision factors. Meanwhile, the domestic researches on deciding a transshipment port are as follows: a study of inducement strategies of transshipment cargo (Bae Byeong-Tae, 1999; Jeong Tae-Won and Gwak Gyu-Seok, 2002; Park Yeong-Tae and Kim Byeong-In, 2003), a study of transshipment port decision based on ISM and AHP technique from the viewpoint of a global container shipping company (Baek In-Heum, 2007), and a study of selection attribution for transshipment port from the viewpoint of a shipping company at home and abroad (Park Byeong-In and Seong Suk-Gyeong, 2008).

However, these previous studies are mainly focusing on the inducement strategies for transshipment cargo or are suggesting the selection attribution for transshipment port as well as the method to select key attributes. But these previous studies have shown that according to the questionnaire respondents (such as a shipping company, cargo owner, importer and exporter, forwarder), their study results are different. In this respect, it means lack of consistency and validity. Moreover, these preceding studies have a limitation in the sense that they are trying to find decision factors only by way of questionnaire, not performing an analysis based on actual data.

Therefore, differently from those preceding studies, this study has analyzed 10-year actual data for comparison analysis. That is, by using Logit regression model and the regression model of Veldman (2008), this study has performed a quantitative analysis in order to suggest selection factors for transshipment port, so that it can be a more practical research.

Comparison of the results earlier Veldman's research concerning the Northwest European market shows that the outcomes correspond rather well in terms of the resulting choice or demand elasticity. This paper shows that the use of Logit Models with respect to transshipment port choice leads to useful findings for port planning. This research in combination with earlier research by one of the authors for transshipment port choice in Northwest Europe is a step forward in the field of transshipment port choice.

3. PORT CHOICE MODEL OF TRANSHIPMENT CARGO

3.1 factors for transshipment port choice

Prior to suggesting port choice model, the factors are to be selected from experts who are responsible to design the shipping liner route. The 15 items to be surveyed are collected from the published papers. The collected items are questioned for measuring importance degree with 5 scores from senior managers of major container shipping liners. The way of collecting questionnaire has been performed by direct interview with a responsible person in a shipping company, or visiting in explanation of the purpose in front of a group of responsible officers in shipping companies. The questionnaires are sent to 30 container shipping liners including both domestic and foreign companies. The sequence of priority is listed table 1. Among these factors, qualitative service factors such as container handling capability, berth facility capability, feeder frequency, feeder network, free time, overtime storage fee etc., can be represented as a proxy variable.

Table 1. Score of transshipment port deciding factors

Item	Response	Mean	SD
Cargo Handling Capability	17	4.294	0.686
THC	17	4.176	0.809
Berth Facility	17	4.059	0.748
Feeder Frequency	16	4.000	0.632
Main Route Location	16	4.000	0.632
CY Facility	17	3.941	0.748
Feeder Network	17	3.941	0.966
Cargo Volume	15	3.933	0.961
Free Time	17	3.706	0.985
Port Dues	17	3.471	1.007
Overtime Storage Fee	17	3.412	0.87
Incentive	17	3.294	0.92
CIQ Service	17	3.294	0.985
Providing Berthing Priority	16	2.938	0.929
Bunker Supply, Ship Repair Service	17	2.824	1.074

In order to make independent variables, cost factor, service factor which is composed of local cargo as attractiveness, total cargo volume as mohring effect, and deviation cost factor and incentive factor can be drawn .

3.2 Model specification

The probability that a shipping company in region (r) select transshipment port (p) can be expressed as:

$$P_P^r = \frac{e^{U_p^r}}{\sum_{p=1}^{p=P} e^{U_p^r}}, (P = 1, 2, ..., P)$$

U is the "utility" attached to transshipment port (p) by shipping liner in region (r) and p the index of the transshipment port in a total of P ports.

Considering Veldman's model (2008) and the factors to be surveyed, the utility function is modified as:

$$U_P^r = \alpha_1 CT_{pr} + \alpha_2 CI_{pr} + \alpha_3 CD_{pr} + \alpha_4 L_{pr} + \alpha_5 M_{pr}$$

where *CTpr* is the sum of feerdering cost *CFpr* and mother ship access cost *CMpr*. The feeder transport cost *CF* is incurred between transshipment port and feeder port (*p,r*) in r region; *CM* is the mainline access cost to transshipment port; *CI* is the incentive between transshipment port and competition port (*p,p'r*); *CDpr* is deviation cost between transshipment hub port and feeder port. *Lpr* represents the attraction of a port given its volume of local cargoes. *Mpr* represents the total handling throughput of a port including local and transshipment cargo. This is a part of Mohring-effects (Mohring, H.,1972) and expressed as a function of the level of port throughput. As feeder calling frequencies increase, wait times of cargo decrease, demand increases, and transit frequencies can increase again. This effect can be used as substation variable of port service. The Greek symbols α1, α2, α3, α4 and α5 are the coefficients of the utility function.

By taking for each region (r) the ratio of the market share of transshipment port (p) and of an arbitrarily chosen basic port (p*), it follows from (1):

$$\left(\frac{P_P^r}{P_{P^*}^r}\right) = \frac{e^{U_P^r}}{e^{U_{P^*}^r}}, = e^{U_P^r - U_{P^*}^r}$$

Combination of equations (2) and (3) and taking of logarithms leads to:

$$Ln\left(\frac{P_P^r}{P_{P^*}^r}\right) = \alpha_1(CT_{pr} - CT_{p^*r}) + \alpha_2(CI_p - CI_{p^*}) + \alpha_3(CD_p - CD_{p^*}) + \alpha_4(L_p - L_{p^*}) + \alpha_5(M_p - M_p^{\ *})$$

3.3 Variable Description

Dependent Variable

$Ln\left(\frac{P_P^r}{P_{P^*}^r}\right)$ is the share of transshipment in the port of Busan among total transshipment in the region.

Independent variables

Selecting independent variables is dependent on research outputs on the topic. Researchers insist that deciding transshipment port is influenced by cost, location, service factors like productivity and incentive system etc.

CTpr-CTpr' is the total cost difference between the port of Busan and Shanghai port for moving containers from origin to destination in North east region This cost is composed of operation cost, running cost and logistics cost.

CIpr-CIp' is the incentive difference between the port of Busan and the port of Shanghai, where THC of deep sea volume is discounted with some percentage or where compensation for the growth of transshipment compared with a previous year throughput. is paid to shipping company.

CDpr-CDp' means the difference of deviation cost from main line route to the port of Busan or the port of Shanghai. In Northeast Asia, Traditionally, main trunk route towards USA is established via Singapore, Hon Kong, Kaosiung, Busan and Yokohama to Los Angeles (Figure 2).

Lp-Lp' can be obtained by the ratio of local cargo of Busan and the region. This is proxy variable representing attraction effect.

Mpr-Mp' can be obtained by the ratio of total handling cargo of Busan and the region. This is proxy variable representing mohring effect.

4. THE RESULT OF MODEL TEST

The multiple regression model is tested with 10 years data. The authors selected one observation per year, therefore the number of data is 10. The reason why just 10 data is selected is caused by the attribute of dependent variable and independent variables. If we increase the number of observation by quarterly or monthly, the result of analysis shows the anomaly, i.e. the significant probability of most of independent variables is under 5 %.

Table 2. Variable definition of model

Variable		Definition	Unit
Dependent Variable	Y	Ratio of Transshipment of Busan with Region Transshipment	Ratio
Independent Variable	X1	Ratio of Local Container of Busan with Region Local Container	Ratio
	X2	Difference of Incentive Payment of Transshipment	US $
	X3	Difference of Mother Vessel Deviation Cost of Transshipment	US $
	X4	Total Transshipment Cost of Mother and Feeder	US $
	X5	Ratio of Total Handling Container of Busan with Region Total Handling Container r	Ratio

4.1 Step 1 Model Test

The regression model has been tested in 3 steps. In the beginning, 5 variables are selected as independent variables similar as Veldman's Model (Veldman 2008). The result shows that X1(mohring effect) and X4(total cost) are accepted, and X2(deviation cost) and X3(incentive payment) is rejected under 5% significance level. As the adjusted R square is 0.986, this

The coefficient of regression model is as follows.

Variable	Non Standard Coefficient		t	Significance Probability	Muliticollinearity	
	B	S.E			Tolerance	VIF Variance Inflation Factor
A	0.1750000	0.0190	8.9940	0.000		
X1	0.0200000	0.0970	0.2020	0.841	0.0060	175.917
X2	-0.00000214	0.0000	-1.6820	0.102	0.3590	2.7830
X3	0.00000412	0.0000	0.7040	0.486	0.0510	19.6870
X4	-0.0000043	0.0000	-11.4980	0.000***	0.1970	5.0730
X5	0.7830000	0.1180	606440	0.000***	0.0040	247.471

* p<.1, ** p<.05, *** p<.01

4.2 Step 2 Model Test

According to step 1 test, the model to be tested is modified with deletion of 3 variables which show weak significant probability. The independent variables to be selected are X4(Total Cost) and X5(Mohring Effect). The result of test is that adjusted R square is 0.995. According to ANOVA, significance probability is 0 and this indicate effective model. Furthermore, the fact that correlation indicator, VIF(variance Inflation Factor), is less than 10 means any dependence does not exist in between independent variables. The coefficient of model is Y = 0.179291199 - 0.000004782589 X4 + 0.810489669X5. This expression will be used for sensitivity analysis.

Variable	Non Standard Coefficient		t	Significance Probability	Muliticollinearity	
	B	S.E			Tolerance	VIF Variance Inflation Factor
A	0.17000000	0.0050	36.597	0.000***		
X4	-0.000004503	0.0000	-21.695	0.000***	0.629	1.590
X5	0.82200000	0.0090	87.179	0.000***	0.629	1.590

* p<.1, ** p<.05, *** p<.01

4.3 Step 3 Model Test

According to step 2 test, the model is to be expanded to identify which one has stronger impact to attract TS cargo between mother ship cost and feeder ship cost. Keeping X5 in the same, X4 independent variable(Total Cost), is required to expand into X6 and X7. X6 means total cost difference due to mother ship operation from Hongkong to Busan, while X7 means total cost due to feeder ship operation from Dalian to Busan. The result of test is that adjusted R square is 0.999. According to ANOVA, significance probability is 0 and this indicate effective model. Furthermore, the fact that correlation indicator, VIF(variance Inflation Factor), is less than 10 means any dependence does not exist in between independent variables. The coefficient of model is Y = 0.94 – 0.9 X5 – 0.00000231 X6 – 0.00001104. The implication from test is feeder ship impacts stronger power rather than mother ship to attract TS cargo.

Variable	Non Standard Coefficient		t	Significance Probability	Muliticollinearity	
	B	S.E			Tolerance	VIF Variance Inflation Factor
A	0.0940000	0.0100	9.7760	0.000		
X5	0.9000000	0.0110	81.880	0.000***	0.163	6.135
X6	-0.00000231	0.0000	-7.9150	0.000***	0.263	3.806
X7	-0.00001104	0.0000	-13.8540	0.000***	0.088	11.415

* p<.1, ** p<.05, *** p<.01

4.4 Port Choice Model for Sensitive Analysis using SD tool

The model which helps estimation of transshipment in Busan can be retransformed to system dynamics model. The SD model is useful tool for presenting social phenomena. The model is composed of total cost and mohring sub system. This system is working with time series and the cost trend, total handling container trend in region and Busan.

In making port choice model using Powersim tool, main level is defined as transshipment ratio of port of Busan. The level variable is linked to inflow variable and outflow variable. In flow variable has sub modules of ‘Morhring Effect’ and ‘Difference of Total Cost’, which are linked to TS inflow.

In a result of SD model, demand curve is drawn like figure 8. For example, if the THC is discounted 10%, transshipment will be increased to 71,364 TEU in Northeast Asia in the year of 2009.

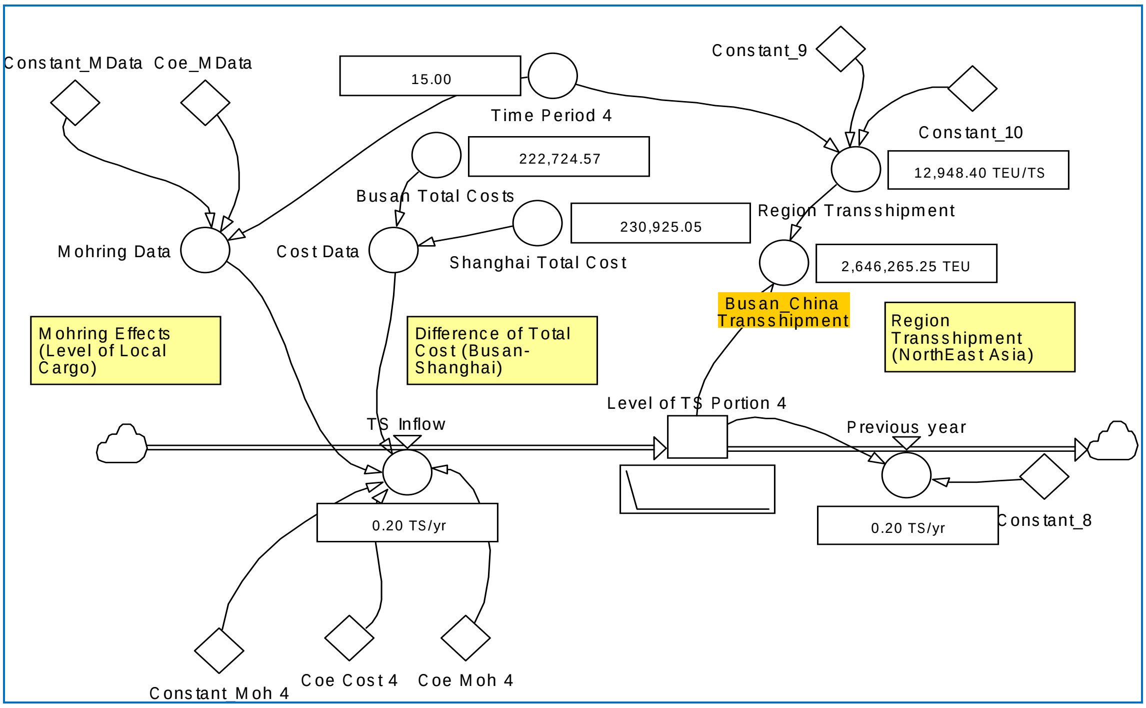


Figure 1. SD Model for Estimation Transshipment

Year	-20%	-15%	-10%	0	10%	15%	20%
2009	144,215	108,533	71,364	0	-69,877	-104,073	-139,755
2010	151,060	113,684	74,751	0	-73,193	-109,013	-146,388
2011	155,433	116,976	76,915	0	-75,312	-112,169	-150,626
2012	156,971	118,133	77,676	0	-76,058	-113,279	-152,117
2013	155,311	116,884	76,855	0	-75,253	-112,081	-150,508

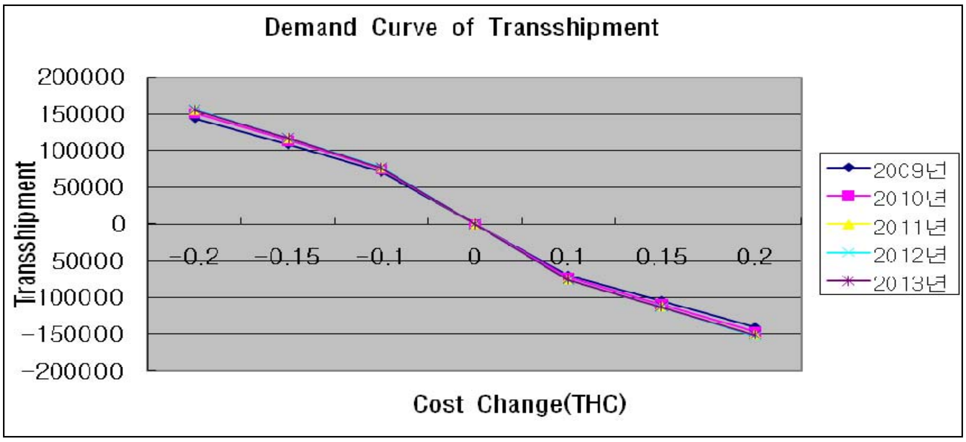


Figure 2. Demand curve of Transshipment

7

5. CONCLUSION

The purpose of this study is to develop port choice model to estimate the trans shipment cargo. The actual data used in the model are based on the 10-year data from 1998 to 2007. Some of these data have been posted by related organizations or the others have been collected during our field visit. The collected data have been used by the above-mentioned selection model in order to calculate the difference values between the two rival ports of Shanghai and Busan. In the first test, the two variables among the five - “the transshipment cargo expenses difference between the port of Busan and the port of Shanghai in both a mother ship and a feeder” and “the rate of Busan Port’s total cargo volume against the regional total cargo volume” - are statistically significant at the significant level of 0.01. Therefore, the second test has been made for these two variables that are statistically significant, and as a result of the second test these two variables again are statistically significant at the level of 0,01. This means that “cost reduction of a mother ship and a feeder ship” and “total cargo increase at the local port” are significant factors for transshipment cargo volume. The third test has been conducted after the mother ship’s expenses and feeder’s expenses have been separated, and the results of this test have revealed that these two are statistically significant at the level of 0.01, and that the feeder’s expenses carry more significance than the mother ship’s expenses. This means that more incentives should be given to the feeders which are suffering financial difficulty.
Based on the quantitative analysis, port choice model to estimate the transshipment cargo is developed using PowerSim tool. The system is useful to anticipate the future volume in adjusting independent variables like “the transshipment cargo expenses difference” or “the rate of Busan Port’s total cargo volume against the regional total cargo volume.”

ACLNOWLEDGEMENT

"This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency" (N I PA-2009-C1090-0902-0004)

6. REFERENCE

Bae, Byungtae (1999). The Devices to Increase Transshipment Cargoes of the Port of Pusan. *Journal of Korea Port Economic* Association, *15*: 179-208.

Containerisation International (2007). *Year Book 2007*.

Drewry (2005), *Ship Operation Costs Annual Review and Forecast, June 2005.*

Drewry(2006), *Annual Container Market Review and Forecast 2006/2007.*

Hiroshi Ohashia, Tae-Seung Kimb, Tae Hoon Oumc, Chunyan Yuc (2005). Choice of air cargo transshipment airport: an application to air cargo traffic to/from Northeast Asia. *Journal of Air Transport Management*, *11*.

Inhum, Baek (2005). Transshipment Port Selection ISM&AHP. *The Journal of Shipping and Logistics*, *54(June)* : 43-64.

Jung, Taewon. and Kyusuk, Kwak (2002). Strategies to Attract Transshipment Container Cargos from/to China by Korea Ports. *Journal of Transportation Research Society of Korea, 20(2)*: 7-16.

Lirn, T. C., H. A. Thanopoulou and A. K. C. Beresford (2003). Transshipment Port Selection and Decision-making Behavior: Analysing the Taiwanese Case. *International Journal of Logistics: Research and Application, 6(4)*: 229-244.

Lirn, T. C., H. A. Thanopoulou and A. K. C. Beresford (2004). An Application of AHP on Transshipment Port Selection: A Global Perspective. *Maritime Economics & Logistics, 6(1)*: 70-91.

Mohring, H. (1972). Optimization and Scale Economies in Urban Bus Transportation. *American Economic Review,* 591-604.

Ng, K. Y. A. (2006). Assessing the Attractiveness of Ports in the North European Container Transshipment Market: An Agenda for Future Research in Port Competition. *Maritime Economics & Logistics, 8*: 234-250.

Ocean Shipping Consultants (2007). *Container Port Strategy, Emerging issues.*

Ocean Shipping Consultants (2007). *Marketing of Container Terminals, 2004.*

Park, Byungin and Soukkyung, Sung (2008). The Decision Criteria on the Transshipment Container Ports. *Journal of Korea Port Economic Association, 24(1)*: 41-60.

Park, Youngtae and Youngmin, Kim (2003). A Study on the Strategies to attract Transshipment Container Cargoes in Korea. *Journal of Logistics Association, 13(1)*: 95-121.

Vvan Reeven, Peran (2008). Subsidisation of Urban Public Transport and the Mohring Effect. *Journal of Transport Economics and Policy*, 42(2).

Veldman S. and Bückmann E. (2003). A Model on Container Port Competition, An application for the West Europan Container Hub-ports. *Maritime Economics & Logistics, 5(2)*, March.

Veldman S., Bückmann E. and Saitua R. (2005). River Depth and Container Port Market Shares: The Impact of Deepening the Scheldt River on the West European Container Hub-Port. Market Shares. *Maritime Economics & Logistics, 7(4).*

Veldman, S., and Vroomen B. (2007). A model of container port competition: an application for the transshipment market of the Mediterranean, *IAME Conference Athens 2007.*

FISHING PORT IN THE RIAU PROVINCE

Erpiani Siregar[1], Reni Meivayana[2], Sumarto[3], Nerli Haryani B[4], Jonny Zain[5], Mashur Siregar[6]

[1,2,3]Department of Fisheries Technology Product
Faculty of Fishery and Marine Science. University of Riau, Pekanbaru, 28293, INDONESIA
e-mail: egha_12@yahoo.co.id , meivayana_87@yahoo.com, sumarto_thpunri@yahoo.co.id

[4,5]Department of Fisheries Resource Utilization
Faculty of Fishery and Marine Science. University of Riau, Pekanbaru, 28293, INDONESIA
e-mail: nerli_haryani@yahoo.com , jonnyzain_@yahoo.co.id

[6]Department of Marine Science
Faculty of Fishery and Marine Science. University of Riau, Pekanbaru, 28293, INDONESIA
e-mail: mashur_sir@yahoo.com

Abstract: Geographically, Riau Province is strategic enough to be developed, one of the field in fisheries, since 71% of its territory is ocean. Fishing port is one of the facilities in the fisheries sector that need to be developed. It needs to be done for good planning in the management and development. In Riau province there are fishing ports that spread in several districts, where the management system coordinated by the government and the private sector. To find out how to optimize the management of fishery port in Riau province, to do research on fishery port in terms of management systems, facilities, activities, and its legal basis. The method used is literature study method. Results from this study obtained sufficient Fishery Port optimal development and operation of fish landings Dumai Base. But in terms of legal aspects of this port is still much to be addressed, as well as other problems. While other fishing port when viewed from all aspects is still far from expectations. Need to be improved in several ways such as in terms of port planning, facilities available, to the clarity of legal basis. From the results of this study is expected that solutions could offer further consideration for all parties, especially government institutions and actors fisheries for economic improvement through optimization of the management of fishery port.

1. INTRODUCTION

Riau province is one of the provinces in Indonesia, whose territory includes the eastern part of Sumatra Island. And is the Islands Region which stretches from the Strait of Malacca, South China Sea and the strait idol. 2,329,867 total area, 61 km, where 71% of which is ocean. Based on this region, the fisheries sub-sector has an important role in Riau. In geographic Riau region is also at the international trade routes that are strategic to be developed and improved through optimizing the management of fishing ports as the economic base.

Fishing port is one of the facilities in the field of fisheries sector that need to be well planned development, so it can be a great chance in the opening of employment relations and the expansion of business field. Therefore the role of the port must be implemented in the best possible advantage for the nation and state, and may place particular and Indonesia in Riau is generally in a place of honor, and is calculated in world trade.

According Murdiyanto (2000) fishing port is the main business activities of the fishing industry must ensure the success of fisheries business activity, especially fishing in the sea , so the fishery port acts as a connecting terminal operations at sea and on land into a business system in a compact and powerful high order. Further investment in infrastructure by government is expected to be followed by the community. While the opinion Zain (2009) Fishing Port has a strategic role in the development of fisheries and maritime affairs, ie as a center or central activity of sea fisheries. Fishing Port is the liaison between the fishermen with catches users, whether direct or indirect users such as traders, mills, restaurants, etc., is also the place of his various interests interacting coastal communities located around the port.

In Riau Province have fishing port that spread in several districts. Like the fish landing there Dumai type D which is managed by local government. In Rokan Hilir district there Fishery Harbor Island Bangliau sunbathe and some managed by private parties such as those found in the district of Pasir Limau Kapas, especially in the Village Panipahan and Pulau Halang. In Bengkalis regency, there are several bangliau or often called a Pelantar of local residents. According Ariani (2007) Pelantar in the Bengkalis District Bantan particular, there are 8 units that are managed by private parties. And most recently the port of Tanjung Samak Fisheries planned early in 2006 and has started construction of the pier. The process is built in several stages such as system multi years with funds from the state budget and local budget.

To support these functions Fishery Port must be equipped with various facilities which are expected to accommodate all fishing activities. Any port in Riau Province Fisheries has different facilities according to needs and conditions. Possibly due to small budget so that the facilities available are also minimal. From the research Mohanda (2008) regarding the development of facilities at the Pulau jemur Port Rokan Hilir sunbathe, there are facilities in this area include basic facilities, functional facilities, and facilities with very modest conditions and some necessary facilities development and utilization. While the results of research Zain (2009) Improved efficiency on fish landing facilities in the town of Dumai, the facilities available in the form of basic facilities, functional facilities, and supporting facilities is complete, in good condition and very adequate. Improvements in fish landing facilities are also likely to increase from year to year. While bangliau is managed by private parties contained in the Villages Panipahan and Pulau Halang has several makeshift facilities in accordance with the needs of just such basic facilities in a state of simple and functional. So also with the platform located in regency Bengkalis District Bantan. Ariani (2007).

Meanwhile, if the views of party management. Fishing Port in Riau is managed by the government and private parties. Port management system of private parties and government are different. Usually if the bangliau managed by private parties, no contract of employment with the fishermen who use the harbor for fishing activities. As the results obtained should catch landed there, as well as operational needs to be purchased to fish in the manager. Meanwhile, the manager also helped the fishermen catch if it distributes in the tender process or the marketing of fish are not sold out. While the port is managed by the government in marketing it is not. Fishermen's catches are not sold out a personal risk manager is not the responsibility of distributing or buy it.

If viewed from the Revised Law. 22 years in 1999 to become law No. 32 year 2004 on Regional Government increasingly stressed the duties and authority of the Central and Provincial Governments, District. Port as one means of transportation require integrated handling both between central and local government as well as between the regional government. Because the port covers issues which are extensive, specialized (lex specialis) and strategic advantage for the nation and the implementation of port state is set and held in collaboration with the Central Government and Local Government. Lex specialist because the port is limited to the area that serves as a gateway that accommodates ships flying the country.

From these phenomena can we know so far synchronization between central government policy and local government about the port in Riau are still far from expectations. So also with inter-regional policy. Some there are who carry out and some not. That problems due to lack of firmness of the Government and the absence of a solution to the problem that caused the existing regulations only mere formality. It is necessary for renewal of legislation that is expected wherever possible to avoid the development of legislation that does not support populist economic activity and harming national interests. Conversely reform legislation must be enabled to utilize the following business investment contributed effectively to regional development within the framework of sustainable national development and be able to anticipate conditions in the period ahead with a consistent, without ignoring the interests of the community.

From this background, this paper will discuss about the Management, by it under facility, and the activity of fishing port in Riau. Then discusses the problems that form obstacles to their operations. Weaknesses, strengths and weaknesses and the strength factor for the occurrence. Results obtained from the comparison between the Port of discussion is conducted to find out where the fishing port which is optimal management. So did about the legitimacy bangliau managed by private parties so far there is no certainty about the position of legal basis concerning the status..

1.1. Problem formulation

From some of the problems outlined in the background can be formulated main problem will be discussed in this paper outlines concerns:

1. Management
2. Institutional and implementation
3. And the problem which includes strength and weakness
4. Factor the cause of the strengths and weaknesses.

1.2 . Purpose

The research aims to identify existing problems in the fishing port of Riau province, which includes:

1. System management
2. Existing facilities and their utilization as well as fishery activities
3. Implementation of Government Regulation on Fishery Port

from the description of the problem conducted Comparison of Fishery Port in Riau qualitatively to determine the best port in accordance with the provisions of the Government.

2. RESEARCH METHOD

This research was conducted using literature study method. Data obtained from the research that raised the issue that covers all aspects of fishing port in Riau Province. Then analyzed and the result is about the management of fishery ports in Riau Province.

3. RESULTS AND DISCUSSION

3.1. Result

3.1.1. Fish Landing Base Dumai City

Figure 1. gate of fish landing Dumai

Figure 2. Office of fish landing Dumai

Figure 3. Place one fish landing facility

1. Management

Fish Landing Base (FLB) Dumai Port fishery type D is contained in Dumai city and managed by local governments. Fish Landing Base Dumai Construction began in 1998 through the State Budget, Budget of Dumai in Riau Province in 2003 and 2003/2004 Budgets with multi-years system. FLB Dumai began operations in January 2004 with the charging activities serving supplies at sea, fish landing, auction and marketing of catches.

By carrying out operational activities to achieve its function well it was established, the Technical Implementation Unit (TIU) Department of Animal Husbandry, Fisheries and Marine City of Dumai. Where the Head of Unit responsible for coo rdinating an d ov ersees th e g eneral ad ministration an d facilitie s field, field b usiness services, p roblems of fishermen and section of the security and order.

2. Facilities

Existing facilities in th e FLB Du mai in clude basic facilities, fun ctional facilities and su pporting facilities. Prin cipal facility owned by the land / soil, pier / jetty, harbor, navigation channel, liaison and drainage of roads, drainage culvert, and bridge.

Functional facilities consist of the fish market stalls and the building where fish auctions, telephone and radio ORARI, clean water supply facilities, ice factories, electric power and diesel fuel filling stations, laboratory quality coaching and processing of fishery products (L QCPFP) , offices of F LB office, LOC PFP office, and office KP3 (Police), and there are also facilities for fish and ice conveyance of the wheelbarrow.

While supporting facilities consist of the guardhouse or the Civil Service Police Unit, drill wells, toilets, shop and parking place.

3. Activities

Activity in the fish landing Dumai starting from the entry process of fishing boats in the harbor area, the landing of fish, fish auctions, marketing the catch, charging the departure of the ship until the ship out of the harbor.

Meanwhile, order activity overseen by the Unit of the police civil service for the smooth fisheries activities as well as the base's security police responsible for the overall fish landings.

3.1.2. Fishing port of Pulau Jemur in Rokan hilir Regency

Figure 4. Pulau jemur of Fishery Port Jetty

Figure 5. Fish Auction Place

1. Management

Fishing Port of Pulau Jemur is situated ± 22 nautical miles east of the capital district Pasir Limau Kapas. This port started its construction in 2000 by local government budgets through the budget Rokan Hilir. System management is coordinated directly by the local government. Like its organizational structure, ect..

Port development in Pulau Jemur is done because the island is visited by many fishermen to catch fish that are rich in the sea like fish, swimming crab, shrimp, turtles, etc. Due to the many and abundant fishery potential in this port should be built to help the fishermen fishing port as a shelter when bad weather and also to extend the fishing trip. It also can help the fishermen to fish in stock so that the needs of the operational cost is reduced.

2. Facilities

Contained in the port facilities include basic facilities such as land, docks, and swimming docks. Functional facility that is where the fish auction, the supply of fresh water, and electricity networks. While supporting facilities such as office buildings and housing.

3. Activities

Fishing activities conducted in the port of landing and marketing of fisheries only the catch, which immediately catch the entire stock by toke who come from different places.

3.1.3 Fishery Port of Tanjung Samak Regency Meranti

1. Management
Leather cape fishery port was built in 2006. Initial phase of development begins with the construction of quay length of 120 m. port development system using multi-years system, whereby the funds derived from State Expenditure Budget, Regional Exp enditure Budg et Beng kalis reg ency, Riau province and regional bud gets. Port m anagement will be coordinated directly by the government and the District of Bengkalis Regency Meranti and Riau Province. Meranti is a division of Regency Bengkalis Regency in the year 2008.

2. Facilities
The new facility is completed at this port with a quay length of 120 m. While the unfinished facility construction such as where the fish auction in the year 2008 began yesterday with the wall construction. However, before the finished wall has collapsed.

3. Activities
Fisheries activities in th e port h as not h appened because th ere are still many facilitie s th at h ave not b een resolved. Fishermen t o engage in m ediator-mediator which is managed by pri vate partie s. Because of this situation, the government is planning construction of fishing ports in order to accommodate all fishing activities in the same place.

3.1.4. Fish Landing Places / Bangliau in Rokan hilir and Bengkalis Regency

Figure 7. Fish landing sites of bangliau/Pulau Halang

Figure 8. Fish Landing Places /Pelantar in Bantan

1. Management
Fish l anding s ites are m anaged by private parties. M anagement sy stem has not st ructure b ut t he t ies of c ooperation between the managers with the fishermen.

2. Facilities
Facilities co ntained i n t his fo rm o f fish lan dings primary facility co nsists o f th e pier, and land . While functional facilities that are available in the form of diesel fuel needs and freshwater.

3.activities
Fisheries activities in these fish landing sites include the provisions at sea, landing the catch, auctions and marketing the catch, mooring and moored. Howev er if in Ban gliau R okan fish ery downstream p rocessing activ ities such as m aking tersasi, salted fish, and shrimp stew made therein by a simple facility

3.2 Discussion

In ge neral, as an ec onomic d evelopment i nfrastructure, t he p ort ha s a n i mportant r ole as a regional e conomic-wheel drive. History also notes that before the era of rapidly expanding aerospace, port is the starting point for the growth of a region because of the ec onomic center of the port became the base through bo th i ntra-regional trade as well as interstate. This is because port not only serves the transportation services or lean on the ship alone, but more than that provide employment, trade, recreation, ship repair services and a series of activities derivatives resulting from one economic activity to another. Therefore, a port must have adequate facilities are facilities that have a high power function, so as not to disrupt the activities of the port service users

From several studies about fishing port in Riau Province in terms of management systems, facilities, activities, and the legality of the position of the port of Fisheries, can know the real problems that need to find solutions for the optimization of fishery port as the economic base.

One of the parameters to determine optimal management of fisheries is to know the port that the internal factors include the strengths and weaknesses of a port.

3.2.1. Strengths and weaknesses of fish landing Dumai

Viewed from the actual conditions, such as frequency of fish landings, production and supplies to sea in FLB Dumai in general showed a tendency to increase from year to year. So also with the development and utilization of the facilities had been upgraded. Therefore it should be a central FLB Dumai fishery must have an operational basis to facilitate port operations and all have a strong legal foundation. Likewise with fisheries policies that lead to the use of PPI as the sole center of all fishing activities both on land and sea. But so far it has not happened. This is one of the weaknesses of FLB Dumai in addition to several existing strengths such as relatively complete facilities to accommodate the activities of fishermen and fish camps outside the city.

To know the optimal management of a port necessary to know the strengths and weaknesses of the existing real conditions. Some strengths of this FLB Dumai include:

1. Relatively complete existing facilities to accommodate the activities of fishermen and fish camps outside the city.
2. Frequency tends to increase in the number of fish landings, fish production and supplies landed sea.
3. Optimal utilization of facilities by the fishermen and even fishermen from outside the city of Dumai.
4. The increased utilization of facilities such as docks, harbors and ponds compared to 2004.

Meanwhile, the weakness of this FLB Dumai include:

1. Incomplete operational runway
2. Some functions of PPI has not been run well
3. There are still some untapped facilities such as cold storage, and fish market stalls.
4. Utilization rate exceeds the capacity of facilities such as ice plants, and SPBN that need development.

Cause of the strength in the FLB Dumai fisheries policy by the government due to lead to the physical development of existing facilities in the FLB Dumai, optimal utilization of facilities is done by fishermen and even fishermen from outside the city. While the cause of weakness is the lack of government attention on an operational basis as a strong legal landing. Utilization of cold storage not exist since the fish are landed daily sold out.

3.2.2 Strengths and weaknesses of Fishing port of Pulau jemur

The main purpose of Pulau jemur Fishing port construction is as fishermen shelter when bad weather and also to extend the fishing trip. Also, can help fishermen to fish in stock so that the needs of the operational cost incurred is smaller.

If the views of its potential, this jemur Island has a very rich fishery potential, and strategic enough to be developed. Hence local government to build this port to facilitate any of Fisheries fishery activities. Besides, many fishermen from other areas who perform activities such as fishing at the fishing port of North Sumatra Province, so the marketing of the catch area. But so far the use of the facilities available will be some untapped well, such as fish and cold storage of place auction. on the contrary there are also facilities such as docks that need to be developed in a state of development because it is damaged and is not adequate for use. From the information obtained can be known to some power.

This port such as:

1. The number of fishermen that many
2. Protected Port Conditions
3. Marketing of Fishery area

While weaknesses are:

1. State of the severely damaged pier
2. Ice plant malfunction
3. Non-functioning of the auction Fish Place

Factors to be the cause of this is due to the power port of fishermen who conduct activities in this port come from various regions such as North Sumatra Province, especially the District. Shavings. Location close to the port area so that the number of fishermen who could be a lot of opportunity to boost the fisheries sector as well as marketing. While the weakness factor is the pier made of wooden construction are not well maintained. The next factor is the ice plant is not used because catches obtained are sold out. Usually the entire stock by fishing contractors who come from different places so that the ice plant is not used as well as the fish auction. Usually a direct auction conducted on the ship.

3.2.3 Strengths and weaknesses of the Fishing Port of Tanjung Samak

Fishing Port of Tanjung Samak would become Port Modern planned the first in the province of Riau. Where the latter will be equipped with various facilities such as a port facility modern. The first phase of construction is planned to start construction a long the 120-meter quay. But so far the co nstruction of t he pier i s abandoned because of budge t limitations. Therefore, for the completion of this course should get funding back, because if it is not treated quickly abandoned the development would be in vain even be destroyed.

Facilities to support fishing activities are also not yet available. As the size of fishing vessel over 60 GT for the South China Sea region, the installation of the fish processing industry development, surimi, ice factories, cold storage, retail outlets, running water, electricity and workshop. We recommend that the Government should support organizing various other facilities to support well over fishing activities in this area. If the port management of these fish will be in operation, of course, the economy surrounding communities will rise.

But appears in the sub-district Rangsang especially in its capital district of Tanjung Samak has become a l and tender for certain parties to benefit a larger number of projects built from the region. The proof is found in many projects from government and from Bengkalis in Riau province and the center was built so long, has not been used already a lot of damages. One example of the construction of a fence wall of Fish Landing Place which is located on an area 20 hectares, the project is a project from the state budget, local budget and Bengkalis, Riau province, in 2005 and 2007. Not to mention the use of building yet another wall was collapsed buildings in the area is abandoned just like that, when a large tide of this land flooded by sea water.

Viewed from this can know the strengths and weaknesses of such fishing port as: when seen from the number of fishermen is very prospective to be developed in order to boost economy communities that have been doing a lot of fishing activities in t he platform, with t he construction of this port will particularly help th e fishermen. While the disadvantage, Port development project which is a dilemma many certainly not abandoned the planned direction. Factors causing the weakness caused by lack of good planning, so that the government seems not only the home building and conducted a survey in advance, giving the impression that the project made the original so. While the benefits to society do not exist at all. So impressed with the planned project in vain.

3.2.4. Strengths and weaknesses at the Bangliau/Pelantar in Rokan hilir of district and Bengkalis Regency

Bangliau contained in Rokan Hilir and Bengkalis regency is managed by private parties, where the organizational structure tends to be familial, because managers are usually single families so there is no formal organizational structure. Bangliau is a business area for management. So, in management, the manager of bangliau usually employ several employees to help smooth fishing activities. While existing facilities such as docks simple impression is only made of wood, as well as other facilities.

While the activities conducted in bangliau rohil include landing, auction, and processing. Usually produced from this region include shrimp paste, shrimp, dried shrimp, fresh shrimp, and salted fish. Products from these areas or marketed their catch to Bagan Siapi-api, the island of Java, and in exports to Malaysia and Singapore.

When viewed from the decision of the Minister of Transportation number 26 year 1998 concerning the implementation of Port, the sea, the bangliau including into the dock for its own sake. The decision stated that the pier for its own sake is a jetty and other facilities located within the Work Environment and Local Government Area of Interest Environmental harbor from the sea port, built, operated and used for their own interests to support certain activities. Whereas in Article 26 decision of the Minister of Transportation of the Republic of Indonesia number 26 year 1998 that in order to obtain management approval of the pier must attach proof of ties of cooperation, drawing the layout of the location, had evidence of exploitation of land outside the area working environment.

In terms of its legal aspects should bangliau equipped with multiple licenses or certificates to support and facilitate in the operation include: Permit Place of Business, Fisheries Business License, Permit catching fish, Fishing Vessel Permit, and certificates of airworthiness and hiring a fishing boat. But so far there are still many who do not complete bangliau license.

While the fish landings in the Bengkalis District, managed by private parties with the mediator called the local residents. So far, platform management systems in the regency is almost the same as bangliau in the regency of Rokan hilir. Processing activities on the platform but this does not exist. Only the landing and auction only. Likewise with the permission letter has almost the same as bangliau.

Strengths and weaknesses Bangliau / mediator are:

1. Privately owned businesses / individuals
2. Has an extensive marketing network
3. Bangliau /Pelantar has its own fishing fleet

while the disadvantage is:

1. Inadequate facilities
2. Unclear organizational structure
3. Unfavorable condition of the building
4. Administratively less accurate calculation of fishery production.

Factor cause s of stre ngths and wea knesses is the number o f fishermen t hat m any, m any conn ections a nd relationships, as well as his own advantage. While the weakness was a makeshift facility because the budget / capital was minimal, and no structured organization well as individual managers so that they automatically own coordinate system

4. Conclusions and Suggestions

4.1 Conclusion

From th e results o f th is study con cluded t hat in general th e fish ing port in Riau province is still less th an optimal function as a fishing port. Caused some of the following:

1. Aspects of the law is less clear, in ge neral fishery port in Riau province has no clear legal basis and com plete. For example, such as local government policies concerning landings of fish in the place / port that has been determined so far no. So that there are many such examples of illegal fish landings in the FLB Dumai. Therefore, the port functions is not optimal. So also with the legal basis Bangliau managed by private parties, so far their legality is still u nclear. Because of its business license is less complete, as seen from the Minister of Communications Decree No. 26 of article 26 and Law No.31 of 2004 on fisheries said all activities conducted in the Republic of Indonesia must have a business license. By hence the operational bangliau often hampered due to lack of complete business letter.

2. System management is less well structured, such as fish landing base Dumai. FLB is th e port o f Dumai's best compared to other ports. But so far as administrative TIU FLB Dumai has ne ver been established but had to walk or operate. Likewise with other ports.

3. In some ports lack adequate facilities available. And it is not balanced between the facilities with one another. For example some existing facilities are not used at all, partly exceeding capacities, such as the Fishery Harbour Island sunbathe, port facilities as a very high level of facility utilization has been severely damaged and need repair. While Fish Auction Place and Cold storage is never used at all.

4.2 Suggestion

From this problem several solutions suggested for improving the management system Fishery Port. And hopefully will be able to improve the quality of fishery port in Riau Province as a prosperous economic base of society. It is suggested that some solution that is expected to be taken into consideration for all parties, especially governments and actors in order to improve fisheries Fishery port as expected. Suggestions are presented include:

1. Making / renewal of legislation that leads to:
   - Optimizing the role of the port as a fishing port is legal, so hopefully all fishing activities conducted only at the port.

- Prohibition of the existence of illegal fishing port.
2. Fishing Port Management system improvement
3. Local procurement in the fishery harbor considerable potential such as fisheries Rokan hilir of district, which work together to manage the system with private parties such as the fishery port of Belawan in North Sumatra province.
4. Improvement and development of Port facilities and launch activities to enhance fisheries
5. Private party and government cooperation in terms of port.

5. REFERENCES

Anonymous.2010 . Fishing Port of Tanjung Samak http://www.dumaipos.com/v2/berita.php?act=full&id=935&kat=13. Access date January 10, at 15.00 pm

Anonymous. 2010. Leather Cape port development plan. http://www.riaupos.com/berita.php?act=full&id=1830&kat=7. Access Date January 10, at 15.00 pm

Anonymous. 2010. Tanjung Samak http://surat.riau.go.id/index.php?mod=isi&id_news=1512. Access Date January 10, at 15.00 pm

Anonymous, 2010 http: / / surat .riau.go.id / index.php? Mod = content & id_news = 1512. Access date January 10, at 15:00 pm

Ariani, F. Y, 2007. Fish Landing Place of Management Studies (Pelanto) In Sub-district Bantan Bengkalis Riau Province. *Thesis*. Faculty of Fisheries and Marine Science, University of Riau. 76 pp. (Unpublished)

Mohanda, 2008. Studies of Fishing Port Development Pulau Jemur, Rokan Hilir of Regency in Riau Province. *Thesis*. Faculty of Fisheries and Marine Science, University of Riau. 57 pp. (Unpublished).

Murdiyanto, B. 2000. Port Role Vol. 1 The role, functions and facilities. Jakarta. 41 pp.

Pane, P.R.A, 2005. Evaluation of fish landing of Dumai in Riau Province. *Thesis*. Faculty of Fisheries and Marine Science, University of Riau. Pekanbaru. 104 pp. (Unpublished).

Sudibyo. 2010 R. E. Port role. http://www.mappel.mappel.org./rekomendasi-mappel/ port-of-age-region-pre-autonomy. Date of access. February 1, 2010. At 10:00 pm.

Zain. J., 2009. Improved Analysis of Efficacy Dumai fish landing facilities. Riau University. Research Institute. Pekanbaru. 45 pp.

AUTHOR BIOGRAPHIES

Erpiani Siregar is a student in Department of fisheriesTechnology Product, University of Riau, Indonesia. She was born on February, 22 1986 at Batang Toru, Indonesia. Her hobbies are travelling, organiziting and Swimming.

Reni Meivayana is a student in Department of fisheriesTechnology Product, University of Riau, Indonesia. She was born on May, 18 1987 at Pekanbaru, Indonesia. Her mail address is <meivayana_87@yahoo.com>.

Nerli Haryani B is a student in Department of Fisheries Resource Utilization, University of Riau, Indonesia. She was born on December, 28 1986 at Malintang, Indonesia. Her mail address is <nerli_haryani@yahoo.com>.

Mashur Siregar is a student in Department of Marine Science, University of Riau, Indonesia. He was born on May, 31 1987 at Padangsidempuan, Indonesia. His mail address is <mashur_sir@yahoo.com>.

COLLABORATIVE PLANNING: INTERORGANIZATIONAL PLANNING APPROACH IN THE MARITIME TRANSPORT CHAIN

Ralf Elbert[1], Roman Grig[2], Fabian Walter[3]

[1]School VII Economics and Management
Chair for Logistics Services and Transport endowed by DB Schenker
Berlin Institute of Technology
Head of Logistics Division
Head of Chair for Logistics Services and Transport, endowed by DB Schenker
Academic Director of DB Schenker Laboratories
Strasse des 17. Juni 135, Sec. H 94
10623 Berlin, Germany
e-mail: elbert@logistik.tu-berlin.de

[2]School VII Economics and Management
Chair for Logistics Services and Transport endowed by DB Schenker
Berlin Institute of Technology
Research Director
Strasse des 17. Juni 135, Sec. H 94
10623 Berlin, Germany
e-mail: grig@logistik.tu-berlin.de

[3]School VII Economics and Management
Chair for Logistics Services and Transport endowed by DB Schenker
Berlin Institute of Technology
Research Manager Transport Networks
Strasse des 17. Juni 135, Sec. H 94
10623 Berlin, Germany
e-mail: walter@logistik.tu-berlin.de

1. INTRODUCTION

The strong improvements in information and communication systems as well as improved transhipment technologies provide the platform for more efficient planning within interorganizational transport chains. Nevertheless, these technologies do not automatically optimize systems based on routines and behavioral patterns, established over decades.

The maritime transport chain has a heterogeneous structure in nature. The various actors – deep sea carrier, short sea shipper, ocean freight forwarder, port terminals, freight operators, rail companies, hinterland terminals, road, and barge carrier and container depots – can be seen as agents whose inevitable functional role in the logistics chain must be coordinated without interfere with their target self-interest and business concepts (Roorda et al., 2010). These locally acting, economically independent actors that are partially in competition with each other, and operate with different business models (e.g. Carrier's Haulage and Merchant's Haulage) on the market, have to adjust capacities for the purposes of their customers. Neither one integrated service provider is able to plan capacity demands over the entire maritime transport chain, nor a cooperation between actors is sufficient enough to ensure efficient planning along the chain. Logisticians – in theory and practice – have to consider the field of behavioral science to describe and analyse planning problems regarding to involved actors' strategic behavior and social embeddedness, too.

The objective of this paper is to illustrate behavioral aspects and topic models for collaborative planning within interorganizational transport chains. Therefore, this paper analyses changes in behavioral routines and governance structures in the interorganizational maritime transport chain using a case study that depicts mutual projections of expected container transhipment.

2. PLANNING IN SUPPLY CHAINS

Three functions of planning can be distinguished: prognosis and uncertainty reduction, coordination, and control (Staehle et al., 1999). For decrasing business risks the accumulation of information and prognostication of future entries is needed. Basing on incomplete and asymmetric information, alternatives and corrective actions have to be considered by the actors. With the availability of better Information (e.g. shorter inquiry periods) plans have to be revised (Mellis, 2004). A systematic use of uncertainty reduction with planning can lead to economic advantages for the maritime actors (Khandwalla, 1975). The coordination by planning adjusts measures to be taken during planning procedure using the coordination principles sequential planning, parallel planning and planning in groups (Bea, 2005). Due to incomplete information, these plans have to be monitored. This control function has two main tasks: First the planned path is compared with the actual development and second the planned path has to be confronted with currently more favorable paths (Mellis, 2004).

Hierarchical production planning is based upon the development of optimized total models and production planning and control (Kistner et al., 2001). The basic idea is to structure decisions by reference to various characteristics and determining these on different levels instead of simultaneous. Due to the useful connection of different decision levels hierarchical planning is a mixture of simultaneous and successive planning (Schmid, 2009). Top-down or retrograde planning determines global planning aims in form of an overall plan which is substantiated in downstream hierarchical levels (Ehrmann, 2004). Benefits of this planning are convergence of aims and efficiency advantages. In bottom-up or progressive planning, this concept is reverse. The lower hierarchical levels are planning their objectives autonomously. These partial plans are being aggregated to one overall plan (Ehrmann, 2004). An advantage is use of more detailed information from operative levels but does not include convergence with business objectives.

A market oriented planning approach regards to identification of new and attractive markets (Braunschmidt, 2006 and extends the planning, basing on common cost- and historical view, with integration of market conditions and target audience needs. Here especially the systematic and analytical deduction of input sizes for planning is a problem. Thes input sizes are usually projected from historical data.

In contrast to maritime transport chain, strategic planning in industry and trade, e. g. planning approaches in assortment policy (Düssel, 2006) or balanced scorecards (Weber et al., 2000), is widespread implemented. Due to its heterogeneous nature, connecting different actors with differing business models relating to specific views of container transport flows, the maritime transport network has to combine all of these views to one harmonized projection. The following table shows the different planning approaches of actors within the maritime transport chain. They are depending on the position and business model of the various actors.

Table 1. Planning Approaches in Maritime Transport Chain

	container volumes	container transhipment	relation based	sector/ branch based	fleet utilization	others
deep sea carrier	X					
short sea shipper	X					
ocean freight forwarder	X					
port terminals		X				
freight operators			X	X		
rail companies					X	
hinterland terminals		X				
haulier (road)					X	
barge					X	
container depots						X

3. MODELLING THE MARITIME TRANSPORT CHAIN

Concerning the hierarchical planning as planning approach in maritime transport chain, a knock-down model of the chain itself and the involved actors with their specific business models (partial models in hierarchical planning) has to be considered. As already mentioned this model has to reference on behavioral aspects, too.

Figure 1 depicts the maritime transport chain and the environment (macro level) covering the actors. The respective business models of the actors (as partial models) are in mutual interaction for customer service (e.g. transport capacity, price and lead times) and cannot be combined in a total model. However, the different business models also have an effect on the other actors involved and thus form dynamically changing cause-effect relationships depending on actors' reactions and behavior. With a combination of dynamic and explanatory methods, such as agent-based modelling and system dynamics, these dependencies can be captured, but have to be complemented with behavioral aspects.

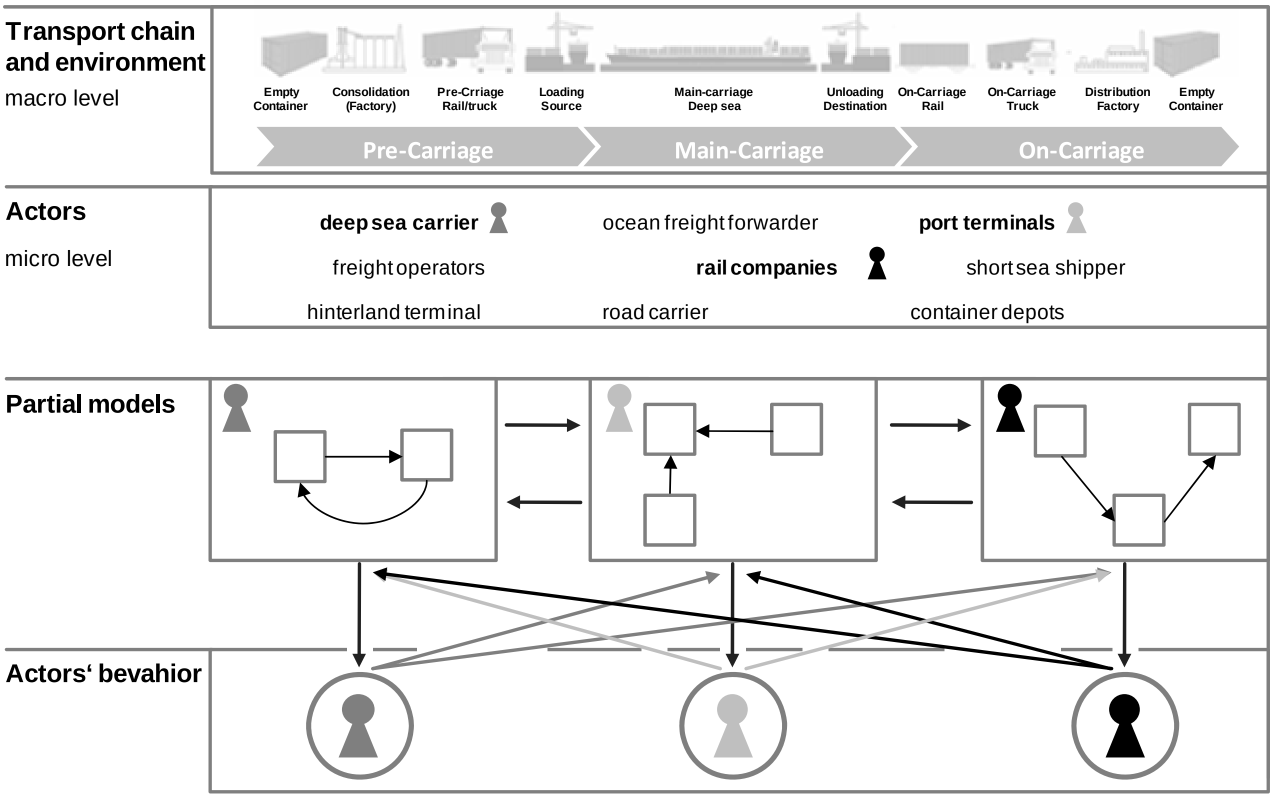


Figure 1. Maritime Transport Chain with Environment

Especially in the maritime transport chain, there are many different modeling approaches and the inconsistency of existing data leads to major problems in integrated forecasting (Jürgens et al., 2008). The SCENES model developed in 2001 is based on the STREAMS traffic simulation model, following the objective of assessing the influence of exogenous factors on the European passenger and freight flows until 2020. The Swedish SAMGODS model is a structured combination of several economic and transport-economic models as a basement for policy decisions about infrastructure projects. The agent-based transport model INTERLOG simulates the interaction between actors in a logistics chain, considering Germany. In general, increased research activities on and implementation of disaggregated models can be determined. Examples are INTERLOG or WiVSim, developed by German DLR Institute of Transport Research (Varschen et al., 2005). These models are based on observations of individual microscopic simulations. Aggregated models, however, are using data, reflecting the actions of specific regional groups (Liedtke, 2006).

As assumption for the planning in logistics and supply chain management – also for interorganizational planning – one behavioral aspect has to be considered in detail: Bounded rationality. This behavioral aspect combined with the fact, that hierarchical planning in maritime transport chain ends at the companies' borders, a new approach in collaborative planning, respecting the behavioral aspects in maritime problem has to be considered. This can be an expert-panel, consisting of representatives of the maritime transport chain. This panel has to be surveyed by a neutral third for collecting expert estimates and aggregating these.

4. CASE STUDY

In the case study expert interviews using the Delphi method and round-table discussions are conducted. Creating projections on future development of container streams within mutual discussions is not only important in phases of static economic growth like estimated in the past years. Integrating experts from all involved actors along the whole transport chain – APL, CMA CGM, Hapag Lloyd, Maersk, TFG Transfracht, Metrans, Polzug, DB Intermodal, DB Schenker Logistics, Deichmann, IKEA, Eurogate Bremen, Eurogate Intermodal, HHLA, Kühne+Nagel, BMW, Bosch, and Siemens Home Appliances Group and DHL – enables projections basing on more different views, experiences and individual behaviors.

For establishing the necessary trust the expert interviews are conducted, anonymized and consolidated by an independent third, here the university. On the one hand quantitative estimations from actor's view could be acquired and on the other occurrences of trends, new aspects and deviations from common behavior of different actors could be identified.

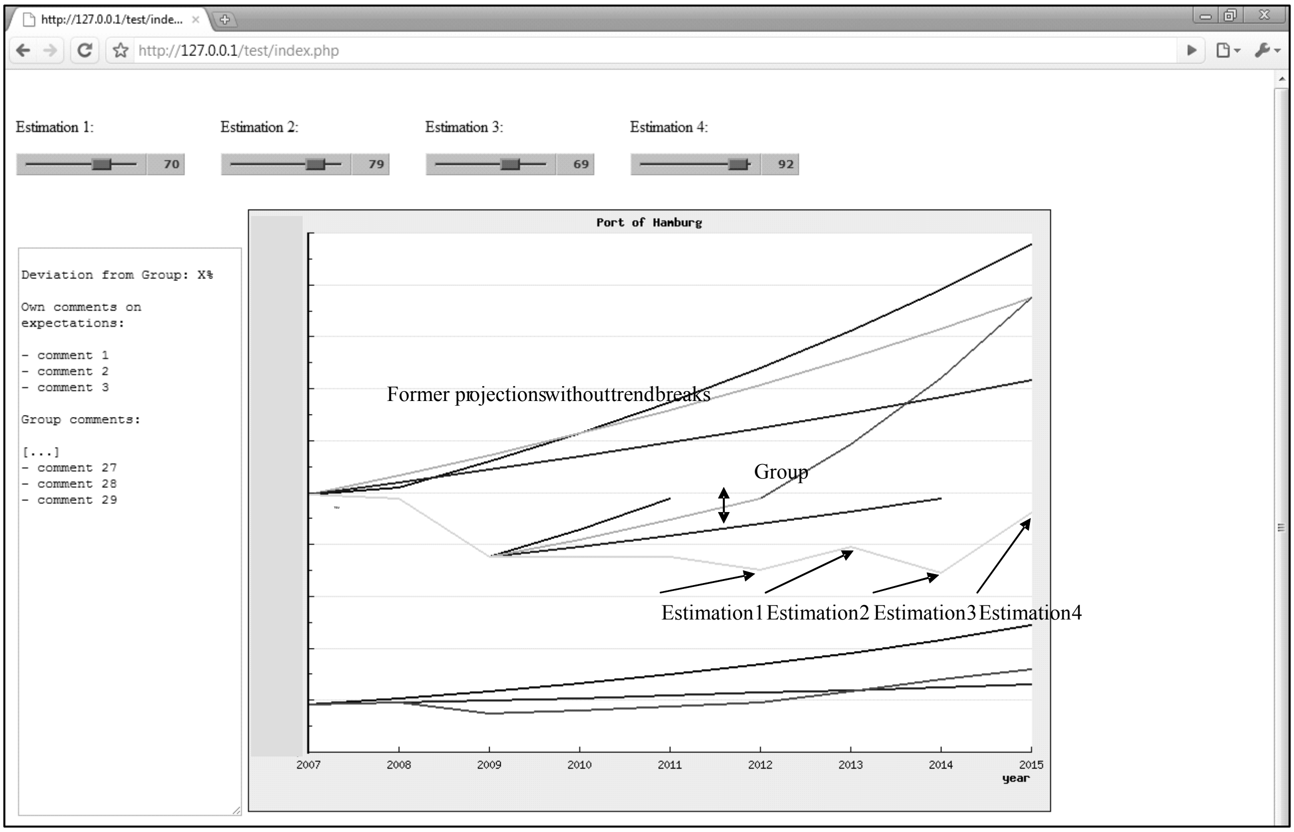


Figure 2. Web-Based Tool Using The Delphi Method (Beta Version)

By interviewing the experts of deep sea shippers, terminals, freight operators and shipping agents in Germany and establishing a multi-client capable survey-tool (see figure 2) the expected results show behavioral changes of the actors in planning by challenging their individual expectations with group expectations and opinions.

In order to identify the relevant quantitative and qualitative estimations of the actors, first a desk research of topic problems and developments within the maritime transport chain was conducted. Basing on these findings an interview guideline for face-to-face expert interviews was derived and sent to the experts in advance. Regarding to the guideline subsequently the interviews had been directly conducted at the experts' place and further open questions and answers were allowed to increase the quality of statements by additional internal information and planning approaches. So the experts' estimations regarding to the seaside and hinterland development in maritime transport chain could be recorded. To establish shorter survey periods the requirements for a first it-based Delphi-prognosis was discussed with the experts. Here especially the planning approaches of the different actors and the willingness for supplying information in such a tool was reviewed.

All interview answers were later on being quantitatively and qualitatively analysed and weighted to the specific market share of the companies. The qualitative analysis showed both commonalities and differences in the experts' opinion on current development of containerized maritime transport.

A quantitative comparison concerning the average growth rate in maritime container traffic first has to unify the different planning approaches. E.g. Carriers are planning in container market volumes while terminals are facing the transhipment of containers (see table 1) including double counts of containers by feeder traffic. If the container is coming back empty to the terminal it is not a double but a quadruple count of one and the same container. After standardization of the container average growth rate estimations by the actors, the datasets were uploaded to a database including all relevant data for the it-based Delphi-prognosis tool.

Although the quantitative expert estimations showed a clearly recognizable positive trend in the transport, the deviation leads to a very early and a quite conservative recovery to the pre-crisis market level in 2008.

For the second round of the survey a link to the it-based Delphi-prognosis tool was sent to the actors. Here the quantitative results were presented and the development paths could be altered by the users with sliders changing the estimated compound annual growth rates (CAGR) for 2010 until 2014.

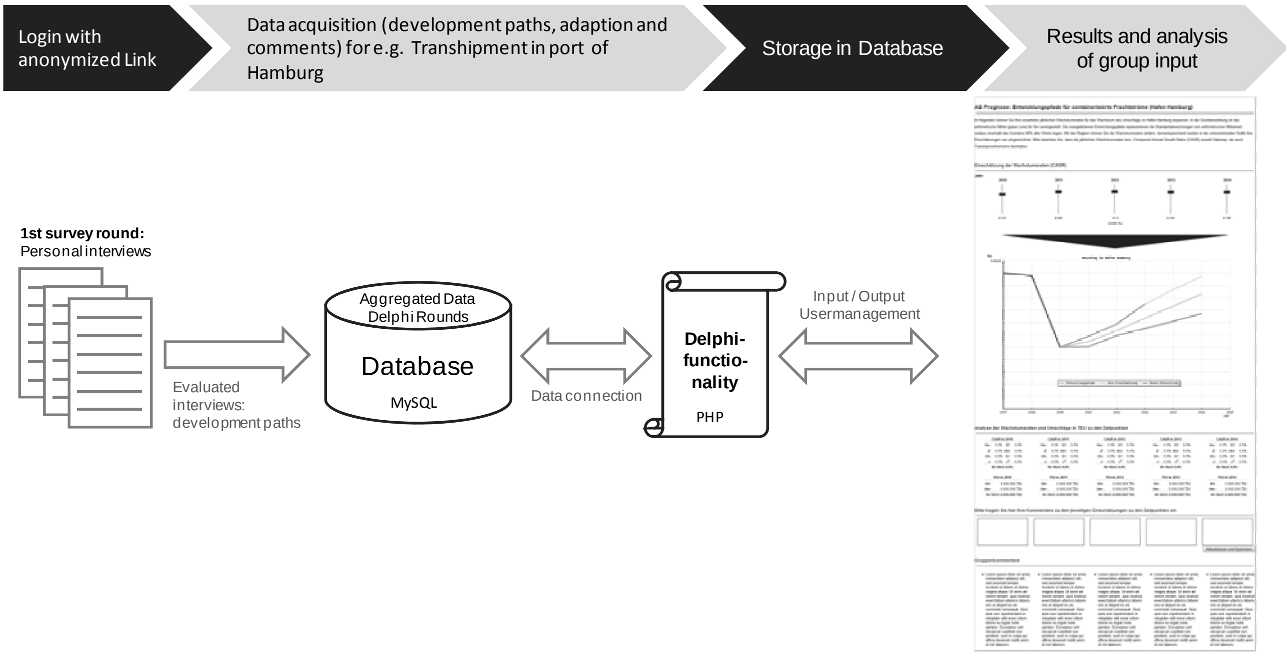


Figure 3. Web-Based Tool Using The Delphi Method (Final Version And Survey Process)

The tool collects the data from the database and computes among other statistical figures the standard deviations from the estimations. Then theses points are presented as two development paths and the average growth rate as path, which can be adjusted by the experts in a second or more round. In the second round the experts varied their expectations with the view on the next quarter. Here the average did not significantly change, so that the first round answers were confirmed.

For the common round table discussion not the quantitative large numbered survey was prioritized but the premises of the actors' planning (planning horizons) and a first it-based Delphi-approach.

Table 2. Actors' Planning Horizons (From Expert-Panel – 15 Interviews)

stagnation gr	owth on a low level	grow with Market	grow stronger than Market	grow twice as Market	no growth information provided
1 2 2 5 3					2

During the round table discussion, the presentation of the common qualitative findings lead to a vital discussion phase, especially when a potential port drift was mentioned. As an actor commented not to notice any shift, the group seized both statements and started a discussion to constitute an overall picture of the statement.

This indicates that even in a closed round with representatives from maritime transport network different views on the container transport have to be collected and directly confronted, to establish a better foundation for a collaborative planning. This furthermore indicates bounded rationality can be circumvented even in a small-numbered survey by using an expert panel. A collaborative planning in the maritime transport chain has economical effects on the market too, e.g. a confirmed positive trend by experts in container volumes has a positive indication for the market growth and thus leads a raise in trade which again increases the container volumes. A confirmed negative trend on the other hand indicates uncertainty and a decrease of market trade.

The discussion about the survey style, whether it should be an empirical sample of a large cross section of maritime actors or a smaller expert panel with relevant actors along the maritime transport chain with a significant market influence, was controversial discussed. As result the relevance of an expert panel instead of an empirical evidence for collaborative planning was confirmed as useful instrument for collaborative planning by the representatives of the maritime transport chain.

5. CONCLUSION

The case study comprises, an expert panel, consisting of representatives of actors across the maritime transport chain, combined with a Delphi-survey tool for shorter planning intervals are a suitable method for a collaborative planning approach. For a better comparison the different planning approaches and horizons have to be unified to one key figure, e.g. whole container volumes or transhipment figures in ports. After unification of the figures, they have to be weighted with the actors' market share in each segment. The quantitative estimations furthermore have to be enhanced by qualitative statements. To eliminate the high impact of behavioural aspects, such as bounded rationality and thus the asymmetric information along the maritime transport chain, the representatives have to be brought together and moderated by a neutral third. Only in direct discussions an overall picture for planning issues can be constituted. This planning approach not only improves the reflection to the actors' own internal planning, but also gives evidence to the market, how the whole maritime transport chain expects the future development in container transport and thus expectations to market growth.

6. REFERENCES

Ahlert, D.; Kenning, P. (2007): Handelsmarketing. Grundlagen der marktorientierten Führung von Handelsbetrieben. Berlin, Heidelberg: Springer-Verlag (Springer-Lehrbuch).

Bea, F. X. (2005): Allgemeine Betriebswirtschaftslehre. 9., neu bearb. und erw. Aufl. Stuttgart: Lucius & Lucius (Grundwissen der ÖkonomikBetriebswirtschaftslehre, 1082).

Berndt, R. (2005): Marketingstrategie und Marketingpolitik. 4., vollst. überarb. und erw. Aufl. Berlin: Springer (Springer-Lehrbuch).

Braunschmidt, I. (2005): Technologieinduzierte Innovationen. Wege des innerbetrieblichen Technologie-Transfers in innovative Anwendungen. Zugl.: Kiel, Univ., Diss., 2004. 1. Aufl. Wiesbaden: Dt. Univ.-Verl. (Betriebswirtschaftslehre für Technologie und Innovation, 49).

Düssel, M. (2006): Handbuch Marketingpraxis. Von der Analyse zur Strategie ; Ausarbeitung der Taktik, Steuerung und Umsetzung in der Praxis. 1. Aufl. Berlin: Cornelsen.

Ehrmann, H. (2004): Marketing-Controlling. 4. Aufl. Ludwigshafen: Kiehl (Modernes Marketing für Studium und Praxis).

Götze, U. (2008): Investitionsrechnung. Modelle und Analysen zur Beurteilung von Investitionsvorhaben. 6., durchgesehene und aktualisierte Auflage. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg (Springer-Lehrbuch).

Grochla, E. (1975): Organisationstheorie. Stuttgart: Poeschel.

Jürgens, S.; Straube, F. 2008. Containerverkehr der Nordrange – Konzeption eines Planungsmodells für den see- und landseitigen Güterverkehr. Internationales Verkehrswesen. April 2008.

Kistner, K.-P.; Steven, M. (2001): Produktionsplanung. Mit 33 Tabellen. 3., vollst. überarb. Aufl. Heidelberg: Physica-Verl. (Physica-Lehrbuch).

Liedtke, G. (2006): An Actor-based Approach to Commodity Transport Modelling, Karlsruher Beiträge zur wirtschaftspolitischen Forschung, Band 20, Nomos, Baden-Baden.

Mellis, W. (2004): Projektmanagement der SW-Entwicklung. Eine umfassende und fundierte Einführung. 1. Aufl. Wiesbaden: Vieweg (Aus dem Bereich IT erfolgreich gestalten).

Roorda, M. J.; Cavalcante, R.; McCabe, S.; Kwan, H. (2010): A conceptual framework for agent-based modeling of logistics services; Transportation Research Part E; pp. 18-31.

Schmid, E. (2009): Koordination im Reverse Logistics. Konzepte und Verfahren für Recyclingnetzwerke. Wiesbaden: Gabler Verlag / GWV Fachverlage GmbH Wiesbaden.

Staehle, W. H.; Conrad, Peter; Sydow, Jörg (1999): Management. Eine verhaltenswissenschaftliche Perspektive. 8. Aufl. /. München: Vahlen (Vahlens Handbücher der Wirtschafts- und Sozialwissenschaften).

Varschen, C. et al. (2005): Ein gekoppeltes Wirtschafts- und Güterverkehrsnachfragemodell unter Verwendung empirischer Daten IN: Nobis, Claudia; Lenz, Barbara, Wirtschaftsverkehr: Alles in Bewegung?, Studien zur Mobilitäts- und Verkehrsforschung, 14, Verlag MetaGIS Infosysteme, Mannheim, S. 193 - 209, Jahrestagung 2005 des AK Verkehr der DGFG, Berlin.

Weber, J.; Schäffer, U.; Ahn, H. (2000): Balanced Scorecard & Controlling. Implementierung - Nutzen für Manager und Controller - Erfahrungen in deutschen Unternehmen. 3., überarb. Aufl. Wiesbaden: Gabler (Advanced Controlling).

Session E1: Simulation

- **Day2: Sep. 16 (Thu.)**

- **Time: 09:00 - 10:20**

- **Chair: Taesik Lee**

- **Room: Camellia, 5F**

COMPARISON OF REAL-TIME CONTAINER STORAGE LOCATION ASSIGNMENT STRATEGIES FOR SEAPORT CONTAINER TRANSSHIPMENT TERMINALS

Matthew E. H. Petering[1*] and Katta Murty[2]

[1]Department of Industrial and Manufacturing Engineering
University of Wisconsin—Milwaukee
EMS E367, P.O. Box 784
Milwaukee, Wisconsin 53201, U.S.A.
e-mail: mattpete@uwm.edu

[2]Department of Industrial and Operations Engineering
University of Michigan—Ann Arbor
Ann Arbor, Michigan 48105, U.S.A.
e-mail: murty@umich.edu

1. INTRODUCTION

Today, almost all overseas shipping of furniture, toys, footwear, clothing, auto parts, electronics components, and computers is done via standardized 20', 40', and 45' long steel containers aboard deep-sea container vessels. As of December 2008, the world cellular fleet consisted of 4661 vessels solely devoted to transporting containerized cargo. The total capacity of these vessels was the equivalent of some 12.1 million twenty foot containers [1]. Four years earlier, in February 2004, the figure was only 6.54 million twenty foot containers [2]. This represents a growth rate exceeding *1% per month*. Such a rapid expansion of the container sector, combined with a heightened concern over customer service and security, has made container shipping a major focus of operational research in the past decade.

In this study, we investigate how the long-run performance of a seaport container terminal as measured in terms of GCR (gross crane rate, quay crane rate) depends on the algorithm that automatically selects storage locations for export containers upon their arrival at the terminal. GCR is defined as the average number of lifts achieved at a terminal per quay crane (QC) working hour and is one of the most important measures of operational performance at a container terminal. To attain a high GCR, the flow of containers back and forth between the shore and the yard has to proceed smoothly, so that QCs do not incur idle time waiting for the vehicles (yard trucks, YTs) that move cargo between the quay and storage yard.

In this study, we evaluate several real-time container storage systems using a fully-integrated, discrete event simulation model that considers the detailed movement of individual containers passing through a pure vessel-to-vessel transshipment terminal over a several week period. We assume that rubber-tired gantry cranes (RTGs, yard cranes, YCs) are used for lifting and stacking containers in the yard and that all equipment is manually controlled by human operators. Experiments consider four container terminal scenarios that are designed to reproduce the microscopic, stochastic, real-time environment at a multiple-berth facility. Results indicate that it is usually more important to choose storage locations that minimize the expected congestion in the yard than to choose storage locations that minimize the expected distance containers travel to and from their storage locations. The experiments do not cover all situations, but nevertheless do provide a meaningful connection between real-time container storage algorithms and the overall, long-run productivity of a seaport container terminal.

2. LITERATURE REVIEW

The literature relevant to this study includes all papers on container terminals that discuss (1) container storage location assignment, (2) simulation modeling, or (3) the literature itself. A total of 63 such papers published in year 2009 or before were found by the authors. Excellent surveys of recent research on container terminal operations include [3], [4], and [5]. A good description of container terminal operations is given in [6]. A concise summary of the various operational decisions made in container terminals is given in [7].

The problem of container storage location assignment is addressed by only a handful of articles in the literature. Of particular interest here is the storage of export containers. Out of 18 such papers, only [8] presents a model in which the arrival, stay, and departure of each container is explicitly modeled. In addition, [8] is the only such article that presents a model that obtains numerical results on *real-time* container storage location assignment. Moreover, no article prior to year 2010 presents a model that shows how alternate real-time container storage location assignment systems affect the overall, long-run performance of a multiple-berth container terminal as measured in terms of GCR or average vessel turnaround time. This research, on the other hand, has produced such a model. This discrete event simulation model is described in [9], [10], and [11]. Out of 40 articles that present container terminal simulation models, these are the only articles that show how proposed terminal designs and/or *real-time* automated decision making systems affect the *long-run*, overall performance (i.e. GCR) of a *fully-integrated*, *multiple-berth* container terminal where the activities associated with individual containers, vessels, QCs, RTGs, YTs, and yard stacks are considered.

3. PROBLEM DEFINITION

The current study has two main goals. The first goal is to evaluate the relative importance of the following four sub-objectives within a real-time container storage management system:

(1) minimizing container travel distance from the quay to the yard during unloading/storage
(2) minimizing container travel distance from the yard to the quay during retrieval/loading
(3) minimizing congestion near container storage location during unloading/storage
(4) minimizing congestion near container storage location during retrieval/loading.

The last two sub-objectives are synonymous with maximizing YC availability in the vicinity of the storage locations during unloading/storage and retrieval/loading respectively. All four sub-objectives are related to the overall objective of maximizing GCR. However, not all sub-objectives can be pursued simultaneously. For example, if the goal is to minimize container travel distance from the quay to the yard during unloading/storage, we would select storage locations that are close to the unloading vessel whenever possible. However, this would probably generate congestion in the vicinity of the storage locations during unloading/storage. Thus, in pursuing sub-objective 1, we would be straying from sub-objective 3. Since not all sub-objectives can be pursued simultaneously, the most important sub-objectives must be identified. Once these are identified, we can derive general rules of thumb that apply regardless of whether a terminal's container storage management system is based on off-line or real-time decision making. Our second goal is to identify one or more specific real-time container storage location assignment systems that strike the optimum balance among the four sub-objectives above.

Our experiments consider four container terminal scenarios that are designed to reproduce the microscopic, stochastic, real-time environment at a multiple-berth facility. Results from a combined 857 million QC lifts worth of simulated activity indicate that it is usually more important to pursue sub-objectives 3 and 4 than to pursue sub-objectives 1 and 2. In addition, the best priority-based storage schemes, which pre-assign to each vessel certain high-priority storage areas where cargo that is loaded onto the vessel is to be stored, perform about as well as the best free storage schemes that do not favor certain areas over others. In other words, mathematical programming has limited relevance within a real-time container storage management system. The experiments are not exhaustive enough to cover all situations, but nevertheless do provide the first direct connection in the literature between real-time container storage algorithms and long-run performance at a multiple-berth seaport container terminal.

4. REFERENCES

[1] *Containerisation International*, February 2009, 19.

[2] *Containerisation International*, March 2004, 5.

[3] R. Stahlbock and S. Voß, "Operations research at container terminals: a literature update", *OR Spectrum* 30, 1-52 (2008).

[4] D. Steenken, S. Voß, and R. Stahlbock, "Container terminal operation and operations research - a classification and literature review," *OR Spectrum* 26, 3-49 (2004).

[5] I.F.A. Vis and R. de Koster, "Transshipment of containers at a container terminal: an overview," *European Journal of Operational Research* 147, 1-16 (2003).

[6] H.O. Günther and K.H. Kim, "Container terminals and terminal operations," *OR Spectrum* 28, 437-445 (2006).

[7] K.G. Murty, J. Liu, Y.W. Wan and R. Linn, "A decision support system for operations in a container terminal," *Decision Support Systems* 39, 309-332 (2005).

[8] R. Dekker, P. Voogd and E. van Asperen, "Advanced methods for container stacking," *OR Spectrum* 28, 563-586 (2006).

[9] M.E.H. Petering and K.G. Murty, "Effect of block length and yard crane deployment systems on overall performance at a seaport container transshipment terminal," *Computers and Operations Research*, 36, 1711-1725 (2009).

[10] M.E.H. Petering, Y. Wu, W. Li, M. Goh, and R. de Souza, "Development and simulation analysis of real-time yard crane control systems for seaport container transshipment terminals," *OR Spectrum*, 31, 801-835 (2009).

[11] M.E.H. Petering, "Effect of block width and storage yard layout on marine container terminal performance," *Transportation Research E,* 45, 591-610 (2009).

THE DEVELOPMENT OF 3D SIMULATION PLATFORM FOR MOBILE HARBOR LOGISTICS AND OPERTATION SYSTEM

Te Xu[1*], Seungsin Lee[2], and Hyowon Suh[1*]

[1] Department of Industrial & Systems Engineering
Korea Advanced Institute of Science and Technology (KAIST)
335 Gwahak-ro (373-1 Guseong-dong), Yuseong-gu
Daejeon, Republic of Korea
e-mail: xute2004@hotmail.com, hw_suh@kaist.ac.kr

[2] WesleyQuest Co.,Ltd
4th Floor Seoul Finance center, 84 Taepyengno1-ga, Jung-gu,
Seoul, Republic of Korea
e-mail: sslee7242@kaist.ac.kr

Abstract: The design concept of Mobile Harbor (MH) is to develop a novel self-travelling harbor system. MH can unload the containers from large tonnage ship outside of the harbor without actually berthing it. Therefore, MH enables receiving large containers ship as well as the other general cargo ship in different styles at small or medium-sized harbors. This new concept of MH brings the new problems such as loading/unloading scheduling, travelling scheduling, berthing scheduling and data management. The approach to the solution of the new problems requires comprehensive and rapid-responsive logistics analysis because of the complexity of the MH system. Simulation method is one of the promising approaches in this case so that this paper proposes a platform of MH system, for integrated comprehensive simulation of MH logistics and operation analysis. It includes 3D resource modeling, 2D & 3D layout design, store planning of resources, performance analysis and database management system.

1. INTRODUCTION

As a traditional freight transport method, the shipping had existed for many years in human being's history, and it still plays a very important role in today's freight transport. There are kinds of ports all over the world, and the requirement of shipping will still increase constantly in future. However, base on the different topography and capacity of ports, the capabilities for various containers ships of these ports are quite different. For some small ports, limited by their topography, it is impossible to allow a large container ship to berth and load/unload the cargoes. Thus, this is the design inspiration of Mobile Harbor (MH), which is a harbor that can be moved out to sea to allow large container ships to unload cargo without having to come into port.

As an innovative maritime logistics system, MH is integrated with various sub systems, including Fast Loading/Unloading System, Floating Body System, Berthing/Mooring System, Quay Cargo System, Container Terminal Resources, and so on. These various and complex components cause the difficulty for testing and evaluating MH operating system, especially in the design phase. Therefore, an efficient and integrated performance evaluation system is very necessary, since these evaluation works are extremely important procedure that cannot be ignored. The improvement of simulation technology these days make it possible to quickly find design error, reduce development time, evaluate performance and guarantee design quality, therefore, it became the popular and frequently applied technology in many industry design area including shipbuilding industry.

Simulation in maritime environments has been developed by the scientists for a long time. In fact both the authors themselves and other researchers had studied technologies devoted to improve performances inside a harbor and on the infrastructures around it (railroads, motorways, etc.) (Briano et al., 2005). However, most of these researches focus on developing mathematics simulation models or 2D process simulation by using ERP and Petri Nets formalism. In 2002, Dong-A University in South Korea devoted to manage Container Terminal Operating Systems using an ERP Approach (Choi H.R. et al., 2002). In the same year, the University of Barcelona in Spain, presented a modeling methodology based on Colored Petri Nets (CPN) formalism to design a software tool useful in port process simulation in order to make easier the design and the management of port terminals (Bacardit et al., 2002). All these researches base on the existing port or the situation that the basic technology or machinery system has been already developed. Therefore, for designing a new port system like MH, the research works are more challenging and difficult.

In this research, by using Siemens NX and Plant Simulation tool, an efficient integrated 3D simulation platform is developed to evaluate the performance of this complex and alternative MH operating system. This platform integrates many technologies like CAD/CAM, Store plan optimization, Machinery & System simulation and Data Base technology. The process for developing this simulation platform is as follows: First, define the 3D CAD resource modeling for designed sub systems, define the relation between systems, find the interaction between various design alternatives, design parameters, and develop a Rule-Based & Knowledge-Based quick responsible system for design changing, finally, base on the information of system components, input data and operating plan and process, the 2D & 3D simulation layout is established.

In there, the information of system components, the input data of simulation layouts and loading/unloading operating plan can be stored in Oracle system of DBMS (Data base Management System). Then, the interface for parameters and analysis function for results can be defined. Finally, after achieving the alternative operating simulation evaluation, by applying the proposed integrated 3D simulation platform for MH, quickly evaluation for reasonable alternatives and design verification can be accomplished.

2. RELATED REASERCH

As mentioned above, there are four main technologies have been applied in this simulation platform: CAD/CAM, Store plan optimization, Machinery & System simulation and Data Base technology.

CAD is one of the computer-aided technologies, which describes the use of computer technology to aid in the design, analysis, and manufacture of products. However, along with the increase of product design complexity and market competition, current traditional CAD tool support is extremely limited for the conceptual design phase. There is need to rapidly conduct design analyses involving multiple disciplines communicating together (trading off such things as performance, cost, reliability, etc.). The solution is to commit more resources at the conceptual design stage to reduce the cycle time by eliminating redesign (Calkins et al., 2000).Therefore, many companies still try to develop their own API (Application Programming Interface) for CAD program (Kim et al., 2005) (Kang et al., 1997). Mean time, a new technology with roots in computer-aided design (CAD) was developed called KBE (Knowledge Based Engineering). Now, many CAD tools have supported the KBE technology, and one of famous application is NX Knowledge-Fusion. The Siemens NX CAD tool provides a knowledge-based editing function called the Knowledge-Fusion. It supplies the capability of rapid redesign base on existing CAD models just by changing the parameters or requirements, not need redraw it. The design concept of MH is to adapt all kinds of container ships, thus, base on the different capacity requirements, the type of MH should be various. NX Knowledge-Fusion technology exactly right satisfies this requirement (Hwang et al., 2007).

Machinery & System simulation are two kinds of virtual simulation technologies. The former focuses on the detail action simulation of single machine, in order to find the design error or machine constraints. The "Virtual Manufacturing" is kind of simulation technology that support the Machinery simulation (Anders et al., 2005) (Chen et al., 2007). The latter choose a factory plant or a product line as the object of simulation to optimize and improve the line performance (Moon et al., 2008) (Choi S.D. et al., 2002). Generally, these two technologies are applied in different researches or different sub tasks for one research. However, more and more simulation tools applied the function that these two technologies can be integrated in one model. In this research, Plant Simulation tool was used to develop the simulation model for MH. The detail action of crane and the whole system operating can be simulated.

Databases remain a critical asset to any enterprise. Although database management systems (DBMSes) are mature, DBMS vendors continue to innovate, extending beyond the traditional relational model by embracing new technologies such as XML, content management, and Web services (Noel et al., 2009). Because MH system has various sub systems to perform the logistics functions, thus, the DB technology was used to store the information data of various components. And if the DB related with cargo cannot be distributed very well for the components, it will cause many problems. Therefore, the DB design of MH operating system is very important. Base on the characteristic or components of MH system, the DB of MH was divided into several sub-databases. This will be introduced in detail later.

In addition, for the evaluation, improvement and automation of container terminal productivity with many components, various existing simulation technologies should be studied (Ramani et al., 1996).

Having all these technology as back ground, the final important job for this platform is to integrate and evaluate the whole system. Therefore, the second development with Plant Simulation tool was achieved, including the input data parameter interface, the result analysis. The details of Store plan optimization is related with optimized solution for programming will be introduced by other papers.

3. MOBILE HARBOR OPERATING PERFORMANCE EVALUATION PROCESS

As a new maritime logistics system, MH system requires internal and external research collaboration with the subject. And the most important task is how these functions are integrated in this platform and how the platform can evaluate the performance of system. After integrating every sub-system base on the subject design, the evaluation of operating performance can be accomplished, showed in Figure 1.

There are four main preparation tasks before the integrated platform established. "Knowledge-Based 3D Parametric Modeling" means the 3D CAD modeling for all sub systems by using NX Knowledge-Fusion technology, here, the interference between each system and collision analysis should be conducted for the integrated design verification. And then, base on the integration rules and engineering knowledge, "System Alternative Layouts Design" can be converted to 2D & 3D simulation model. Once the 2D & 3D simulation models are finished, it can be repeat applied in the different MH simulation scenarios. Then, base on the simulation models, "3D Simulation and Performance Evaluation" and "3D Emulation Assessment and Management Plan" should be accomplished. The latter represents the optimized store plan of containers and MH management plan in port.

After all these processes has been accomplished, an integrated 3D platform can be established, and with the MH management plan, the store plan of containers and the crane operating information, the simulation scenarios can be experimented, and the analysis of simulation result can be promptly gained.

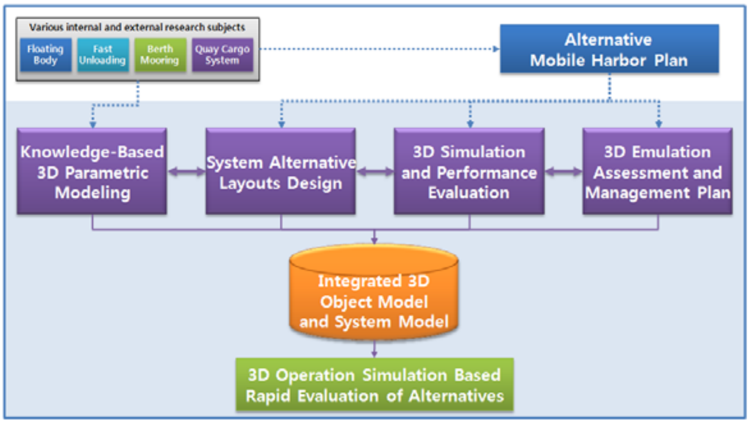


Figure 1. Mobile Harbor Operating Performance Evaluation

4. KNOWLEDGE-BASED INTERGRATED MODELING

4.1 Integrated Modeling and Design Verification Component

The MH is a kind of maritime logistics system, which can load/download the containers from container ship directly, even though in the sea, 2~3m wave striking happens frequently. Therefore, MH includes various sub-systems and components, which are designed to solve the various problems happening in the sea. Through the integration of sub-systems and evaluation of all design, the final design of MH shows in the Figure 2. By using the 3D modeling technology, the interference between systems can be removed, the accuracy and efficiency can be improved, and the productivity can be increased.

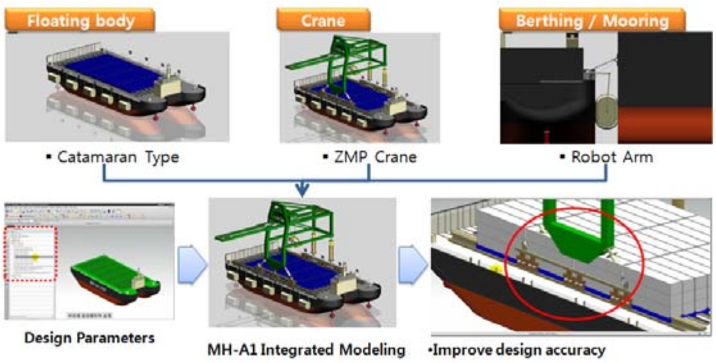


Figure 2. Integrated Design Verification of Mobile Harbor Components

4.2 3D Simulation Library Based on Knowledge-Fusion

As mentioned before, the KBE (Knowledge Based Engineering) was applied in this platform to satisfy the alternative resources. Because design activities are automated, KBE represents a way of capturing design processes. KBE software often have an object oriented structure, where geometry is created and evaluated in terms of classes containing stored rules, and a geometry engine that manages and visualizes the geometry (Sandberg et al., 2008). Chapman and Pinfold claim that KBE can promote organization of the information flow and create rapid design solutions where more ideas and scenarios can be evaluated than without KBE (Chapman et al., 1999). For rapid design of MH and other resources, the knowledge-based design concepts by using NX Knowledge-Fusion technology has been applied. By using the Knowledge-Fusion, the entire 3D model's size (including parts and assembly parts) can be change directly under the requirement of designer for every kind of simulation scenarios. By doing so, cost estimations and warning messages can be generated as a result of design changes. Figure 3 shows an example of this kind of application. One 240TEU MH is directly transferred into a 1280TEU MH, and base on all kinds of constraints and rules, one crane also is added into the MH.

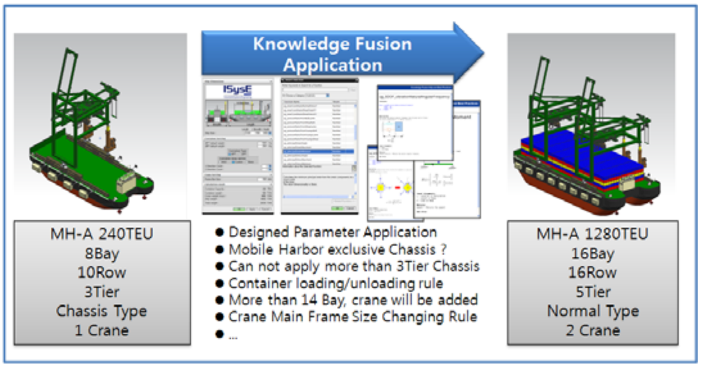


Figure 3. Knowledge-Fusion-based Platform

This function of Knowledge-Fusion is to meet the MH's various constraints and requirements. Base on the customer requirements, business environment and context of the MH, parameters are determined automatically by defining the design rules and knowledge, thus, this system can response the variety very quickly. After the developed model is converted into 3D CAD system, geometric models with details also have been created. Therefore, it can be applied in the design engineering directly. Figure 4 shows the operating interface of this function in detail. The input parameters include the average loading weight, the container types (20FT/40FT), the crane frame size and so on. After inputting the parameters, by using KBE technology, the geometry data can be calculated and transfer into new 3D CAD model.

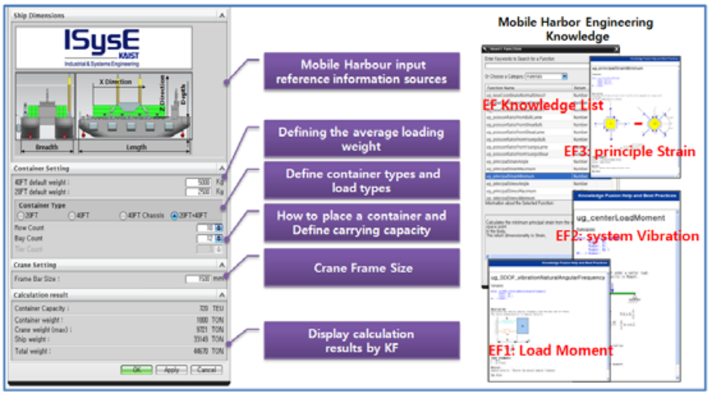


Figure 4. Mobile Harbor Configuration Variation

The construct of MH is showed in Figure 5. One important characteristic of MH design is the container's carrying capacity (Row, Bay, and Tier) which also represents the position of containers in MH. This position is very important for the loading/unloading sequence of the containers. Now, the capacity of standard MH is 240TEU. By using the technology we mentioned, it can be converted into many kinds of MH with different capacity by different requirements.

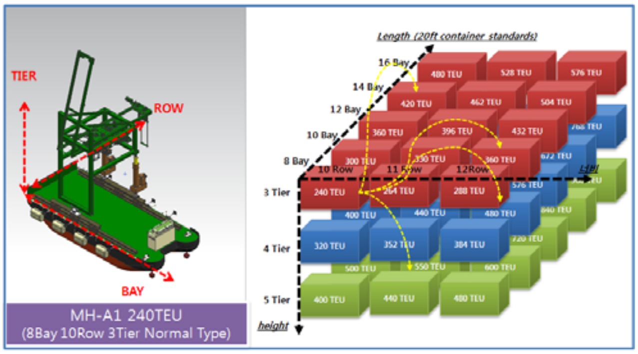


Figure 5. Standard Model Alternative Variation

The capacity (Row, Bay, and Tier) is one of MH construct's constraint. Base on the knowledge of rules and high-speed dedicated unloading chassis (Chassis) availability, the Tier of 240 TEU MH cannot be larger than 3. That means if the Tier is more than 3, it will affect the stability when MH travelling in the sea, and also affect the loading/unloading speed or efficiency of crane. However, along with the increase of loading capacity, the tier of containers may also increase. In addition, if the number of Bay is more than 14, than two cranes are required for one MH.

5. SIMULATION PLATFORM

For choosing the best service system of MH, finding the design errors, evaluating performance and analyzing the enforceability of this system, 3D simulation platform of whole system including the detail action of MH crane have been developed by using Siemens Plant Simulation software. This 3D simulation platform is not only a simple and graphical animation. Many important details of components like cranes, trolleys, containers and vehicles and their mutual relationships are included, and the information data transfer among different sub-systems are also designed. The construct of this simulation platform are showed in Figure 6.

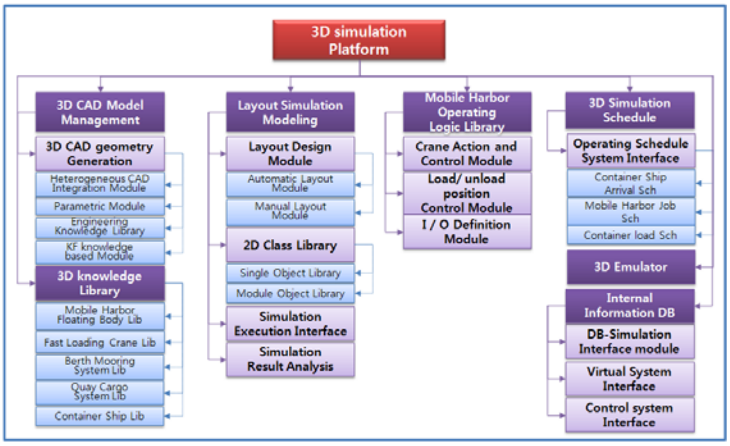


Figure 6. Simulation Platform Components

There are four seven main parts of 3D simulation Platform, and they are 3D CAD Model Management, 3D knowledge Library, Layout Simulation Modeling, MH Operating Logic Library, 3D simulation Interface, 3D Emulator and Internal Information DB. The integration and data transfer between these parts consist of 3D simulation Platform . 3D CAD Model Management and 3D knowledge Library supply the 3D CAD file of every component of MH operating system. Four parts will be explained in detail.

5.1 Operating Logic Library

The operating logic library of MH simulation platform include many kinds of logics, such as the operating offshore implementation and control module of MH crane, the position and sequence logic when containers were loaded/unloaded in the container ship or MH, the scheduling when MH depart from the port and the parking position choosing logic when MH berth around the container ship, and so on. Some of these logics are stored in the Database, and some of them are developed in the simulation models. 3D simulation Schedule supplies all kinds of schedule of MH operation.

On the other hand, the platform supplies the input interface for important parameters. Once these parameters are input by users, the program in the back of platform will connect it with simulation model and database. Through the simulation, all the effect of MH system by these parameters will show in the output data. Figure 7 shows the input data and output data in this simulation platform.

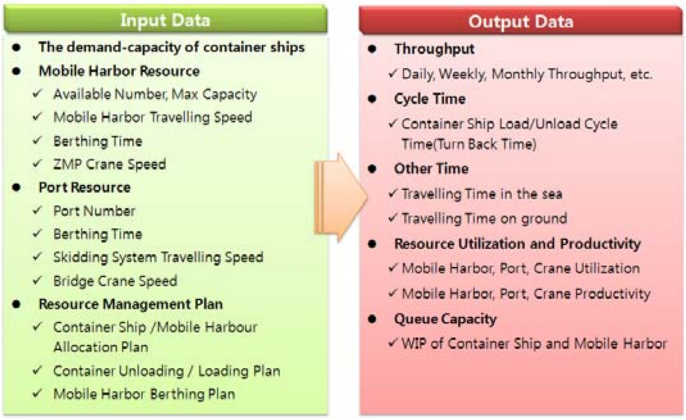


Figure 7. System Input and Output Data for Verification and Comparison

5.2 Internal Information Components

As mentioned above, the database is used in this simulation platform to store much important information. Because there are many kinds of information data, the classification of every data is necessary. Figure 8 shows the construct of the database, and the relationship among every sub-database. The database has been classified five main parts. They are the Information of Platform Users, the Basic Information & Specification of Container Ship and MH, the Information of Simulation Experiments, the Input Information of Container Ship, the Input Information of MH, and the Operating and Task Logs. By the logic rule of MH operating system, the in data in these DBs will be modified frequently to evaluate the performance of system in different situation. Schedule DBs supply the information of container loading/unloading, Specification DBs supply the basic information of resources, and Log DBs record and manage the simulation result information.

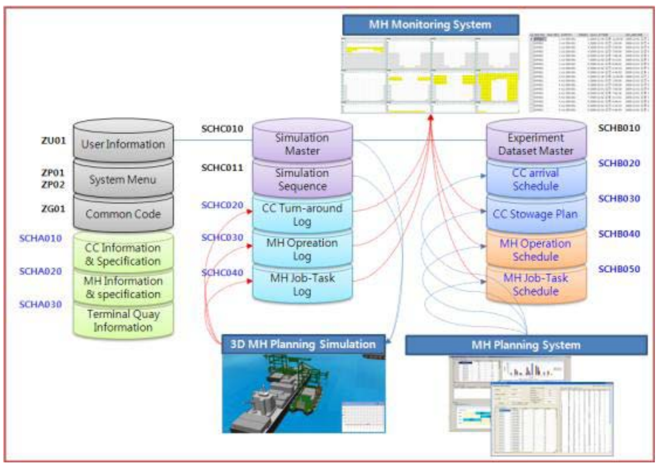


Figure 8. Platform Database Configuration

5.3 Layout Simulation Modeling

Simulation layout in this platform includes three main parts: simulation model, simulation execution interface and result analysis interface, which includes all details of the component size, capacity, speed and can easily change the scenarios base on parameters, design and configuration changes. Among these three parts, the simulation model has two kinds of represented types. One is the 2D layout, and the other is 3D layout. Figure 9 shows the physical and 2D simulation layouts, simulation execution interface and result analysis interface.

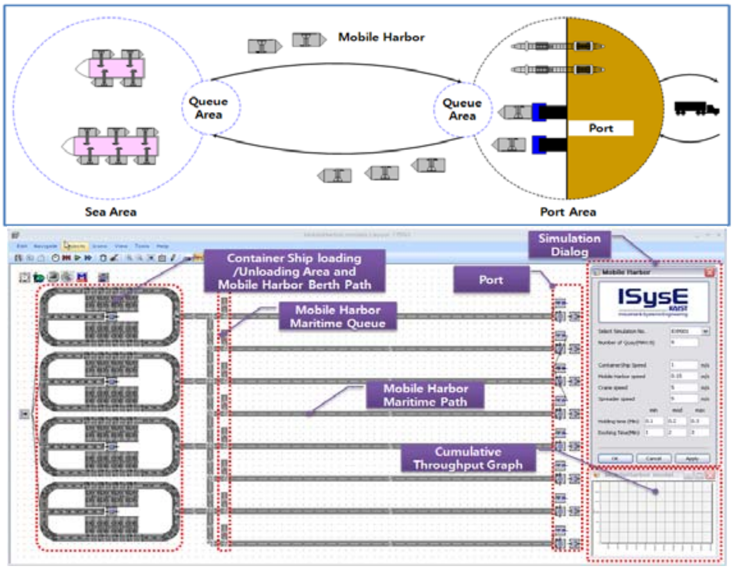


Figure 9. Mobile Harbor 2D Simulation Model

As showed above, the layout includes port area, travelling road in the sea, queue area and the berth area for container ship loading/unloading. By receiving the parameter input in execution interface, the simulation model can simulate the operating of MH system, and all results will show in the result analysis windows. Then the engineering can evaluate the performance of the MH system in different scenarios. This layout also has the more detail type with 3D simulation models, showed in Figure 10. From the 3D simulation model, the details of MH can be represented, especially for the moving action of crane and the containers.

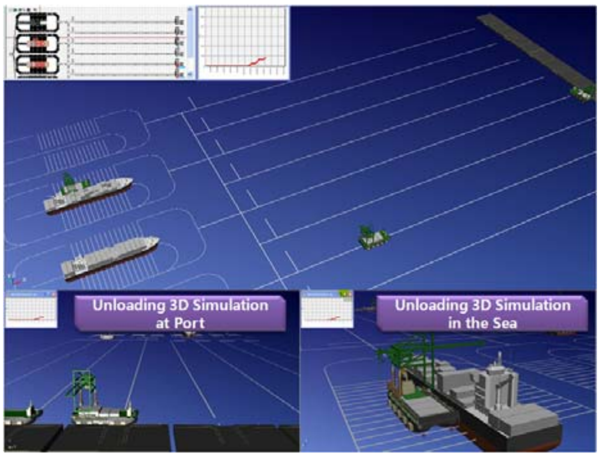


Figure 10. Mobile Harbor 3D simulation Model

5.4 Simulation Result Analysis

Once the simulation of various MH scenarios is achieved, many output data can be gained by the platform. For evaluating the MH performance efficiently, except the basic output data, the result analysis method is also integrated in the 3D simulation Platform. Figure 11 shows an example of result analysis by the platform program. The left chart shows the time phase for every detail task, in there, the details of MH working time can be gained, and the compare between the task plan and simulation result also can be showed by chart. In addition, the compare between different simulation results of same scenarios also can be achieved.

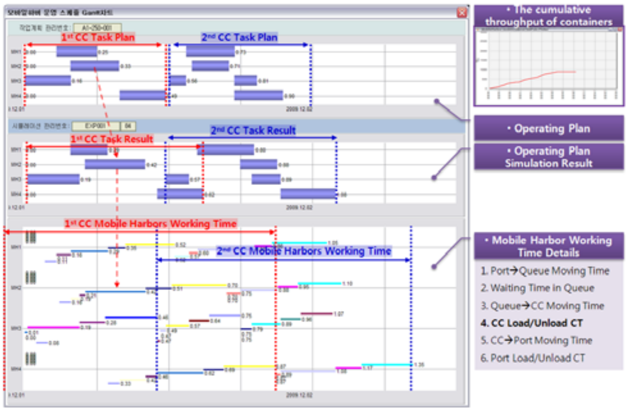


Figure 11. Simulation Result Analysis Example

6. CONCLUSION

In this study, a high complexity and integrated MH System simulation platform was developed. In order to improve the precision and variety of interactions between designs alternatives, design parameters and association, and quickly response to the design constraints and rule change from time to time, a knowledge-based system (3D CAD geometry creation component) has also been developed. Four main technologies have been applied in this simulation platform: CAD/CAM, Store plan optimization, Machinery & System simulation and Data Base technology, which are developed and improved in their own area for many years. Therefore, the work that integrates these mature technologies into one platform is quiet difficult.

First, for building the 3D CAD model library of MH, the design concepts are converted into a 3D CAD systems Geometric Models with details of every the component. By using the technology of Knowledge-Fusion, the 3D CAD models' size and construct can be changed base on the different parameters and requirements.

Then, the information database of whole platform including input data, container ship data, MH data and all kinds of experiments data was established. Once the simulation model execute, the required information data will be gained from this database by kinds of logic program. Therefore, the management of logic for MH operating system is another important task for this simulation platform.

After classifying the logics and establishing the relationship among logics and sub-systems, the simulation layout can be built. Both 2D and 3D simulation models are established in this research. The later is the extend work of former, which represents the many details of MH operating.

Finally, the evaluation part of whole platform was established to test the performance of MH operating system. The evaluation of this platform supports the function of parameter input and results analysis. By changing the parameters of container ship or MH, the performance of MH operating system in different situation can be evaluated.

7. REFERENCES

Anders, J., Wall, J. and Broman, G. (2005). A virtual machine concept for real-time simulation of machine tool dynamics. *International Journal of Machine Tools & Manufacture,* Vol. 45, 795-801.

Bacardit, M., Canudas, J., Piera, M.A., Vila, J., Guasch, A, Amettler, S. and Verges, D. (2002) A Generic Reusable Model of a Harbour System Described at a Macro Level Detail. *Proceedings of HMS 2002,* 3-5, Bergeggi, Italy.

Briano, C., Briano, E.and Bruzzone, A. G. (2005). Models for Support Maritime Logistics: a Case Study for Improving Terminal Planning. *19th European Conference on Modeling and Simulation*.

Calkins, D. E., Nathaniel, E. and Christian, S. (2000). Knowledge-Based Engineering (KBE) Design Methodology at the Undergraduate and Graduate Levels. *Int. J. Engng Ed*, Vol. 16, No. 1, 21-38, Printed in Great Britain.

Chapman, C. B. and Pinfold, M. (1999). Design engineering - a need to rethink the solution using knowledge based engineering. *Knowledge-Based Systems,* Vol. 12, 257-267.

Chen, K. Z., Feng, X. Y., Wang, F. and Feng, X.A. (2007). A virtual manufacturing system for components made of a multiphase perfect material. *Computer-Aided Design,* Vol. 39, 112-124.

Choi, H.R., Park, N.K., Kim, H.S. Lee, S.W., Park, B.J., Kim, J.J. and Lee, K.S. (2002) An ERP approach for Container Terminal Operating Systems. *Proceedings of the 18Th International Port Conference on Increasing Private Sector Involvement in Ports & Transport: An Opportunity or A Threat,* January 27-29, Alexandria, Egypt.

Choi, S.D., Kumar, A.R. and Houshyar, A. (2002). A Simulation Study of an Automotive Foundry Plant Manufacturing Engine Blocks. *Proceedings of the 2002 Winter Simulation Conference,* 1035-1040.

Hwang, I. B. and Lee, S. H. (2007). Development of Knowledge-based Mold Design System using Knowledge Fusion, *Korea Society of Precision Engineering Conference,* 781-782.

Kang, C. S. and Han, S. H. (1997). Parametric Design of a Car Audio Based on Configuration Design Method. *Journal of Korean CAD/CAM Society,* Vol. 2, No. 4, 276-285.

Kim, J. H., Park, J. H., Moon, C. S. and Hwang, Y. K. (2005). A Study on the 3D Injection Mold Design Using Unigraphics API. *Journal of Korean CAD/CAM Society,* Vol. 10, No.6, 381-391.

Moon, D. H., Xu, T. and Shin, W. Y. (2008). A Case Study of Simulation for the Design of Crankshaft Line in an Automotive Engine Shop. *Journal of the Korea Society for Simulation,* Vol.17, No.2, 1-12.

Noel, Y., Gilpin, M. and D'Silva, D. (2009). The Forrester Wave™: Enterprise Database Management Systems, Q2 2009. *For Application Development & Program Management Professionals*.

Ramani, K.V. (1996). An Interactive Simulation Model for the Logistics Planning of Container Operations in Seaports, *Indian Institute of Management,* Vol. 66, No. 5, 291-300, Ahmadabad, India.

Sandberg, M., Johnsson, H. and Larsson, T. (2008). Knowledge-Based Engineering in Construction: The Prefabricated Timber Housing Case. *Journal of Information Technology in Construction,* Vol. 13, 409

CAPACITY AND QUEUEING EVALUATION OF PORT SYSTEMS WITH OFFSHORE CONTAINER UNLOADING

Woo-sung Kim[1], Jong Hoe Kim[2], Hyoyoung Kim[3], Hyeyon Kwon[4], and James R. Morrison[5*]

Department of Industrial and Systems Engineering
KAIST
335 Gwahangno, Yuseong-gu
Daejeon 305-701, Korea
e-mail:{kws0924,windf,hyhouse5,hyeyonnie,james.morrison}@kaist.ac.kr

Abstract: Analytic performance evaluation of port systems that include offshore unloading of containerships, as in Hong Kong's midstream operations and recent mobile harbor concepts, is complicated by the additional container handling step and the finite number of resources available. Commonly used decoupling approximations, that consider each resource as an independent queueing system, may not be sufficient to obtain adequately accurate estimates of system performance. In this paper, we develop several rough cut evaluation techniques for port systems with offshore operations; there are two thrusts. First we focus on container ship service time and capacity analysis. Based on system primitives such as the container capacity of the offshore vessels, the ship unloading speed and the land berth unloading rate, we calculate a lower bound on the minimum number of offshore vessels to serve a given demand for port services. Based on this throughput approach, we in turn obtain an approximation for the ship service. We also develop a Petri net model to obtain more exact results for ship service time. Our second thrust is to develop approximations for queueing time to enter service. Here we use a modified M/G/c queue to model the system. This model is then employed to develop an approximation for the queuing and service time of container ships. With the goal of assessing the accuracy of our approaches, we develop simulations of port systems with offshore components. We discuss the results and highlight which methods are of practical value.

1. INTRODUCTION

As the volume of trade between nations has grown, worldwide maritime container transport has followed. In fact, it was projected in Drewry(2006), that by 2011, the container transportation market demand will exceed capacity. While this projected date has no doubt been delayed by the worldwide economic slowdown, the need for additional port capacity will continue. To address this problem, various countries have increased their container terminal capacity in hopes of attracting the global container market. One of the traditional solutions to increase capacity is to construct new berths. However, this approach may not be ideal because it requires substantial time, is very costly, and is not matched with green development or ecological technology. Most container terminal construction uses large amounts of artificial ground and concrete and consumes large tracts of seashore. An alternative is the mobile harbor system. This system is detailed in Suh N.P (2008) and is intended to address the existing problems of traditional container terminals, such as a lack of berth space and water depth. To efficiently design and operate a mobile harbor system, it is important to evaluate the measures of system performance such as throughput and ship queuing time. However, commonly used queuing approximations do not well fit the mobile harbor system. Therefore, new techniques are required to assess the system performance.

In this paper, we study evaluation techniques for the mobile harbor system. First, conducting throughput analysis, we calculate the minimum number of mobile harbors to serve a ship. With this, we can calculate ship service time under ideal service assumptions. After relaxing the assumptions, we propose another approximation for ship service time. Second, we conduct queuing analysis. Employing an M/M/c queuing model, we can obtain the number of MHs and berths to achieve a required service level. Waiting time to service time ratio (W/S ratio) is used as service level metric. We also develop approximations for the cycle time of a ship. Modeling the system as a modified M/G/c queuing system, we develop approximations for queuing and service time of container ships. To verify our approximations, we develop simulation, and compare with our approximations.

This remainder of the paper is organized as follows. In section 2, we describe the mobile harbor system adopted in this paper. We introduce how the mobile harbor system operates and define variables which we use in this paper. In

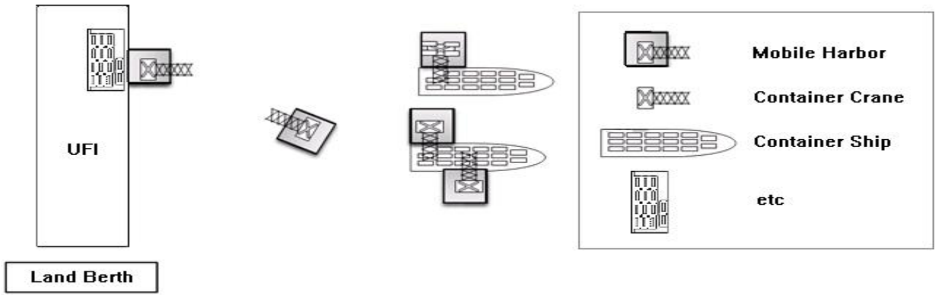


Figure 1. Mobile Harbor (MH) system

section 3, throughput analysis is conducted. With different assumptions, several techniques to evaluate system performance are introduced. Two techniques which calculate a minimum number of offshore vessels to serve ships are introduced under ideal assumptions. Using the techniques, we can estimate the ship service time without waiting time. Then, based on a petri-net model, we conduct throughput analysis. We can obtain throughput and cycle time assuming no queuing. In section 4, approximations for the queuing and service time of container ships are developed. We model the mobile harbor system as a modified M/G/c queuing system. We compare the approximation with simulation results. Concluding remarks will be presented in section 5.

2. SYSTEM DESCRIPTION

In this section, the concept of the mobile harbor system is introduced. In traditional harbor systems, ships dock at land berths to unload containers. There are challenges associated with the traditional system. Due to container transportation market trends, the size of ships has increased (Drewry(2006)). However, most ports do not have the capability to serve the new containerships because of water depth. Insufficiency of berth space is another problem. The mobile harbor system addresses these problems. In the mobile harbor system, mobile ships/barges, each possessing its own container crane and storage capacity for hundreds of containers or more, travel on the ocean to vessels anchored offshore. There it employs its onboard crane to conduct loading and unloading of cargo at sea. We call such mobile crane and container storage ships as mobile harbors (MH). When a container ship arrives to the offshore operation location, mobile harbors provide service to the ship. The mobile harbor next transports the containers to a dedicated land berth. Once arriving to the dedicated land berth, the mobile harbor releases the containers to the land berth. Because the (un)loading operations between a ship and a MH is conducted on the ocean, it can more readily guarantee water depth. At the inland berth, an Ultra Fast Interface (UFI) with high container transfer speed moves containers between the berth and the MHs. Therefore, the MH system can minimize the dependency between the port capacity and the length of the seashore. Figure 1 depicts a MH system.

The following notation is used in this paper:

λ	Arrival rate of a ships (ships/day),
cMH	Capacity of a mobile harbor (TEU/MH),
nMH	Number of mobile harbors,
nB	Number of berths,
nTU	Number of containers in a ship,
nD	Maximum number of MHs that can simultaneously be docked with a containership,
Vt	Travel speed of a Mobile harbor (Km/hour),
Vs	Operation speed of a Mobile harbor at ship (Containers/hour),
Vb	Operation speed of a Mobile harbor at berth (Containers/hour),
d	Distance between offshore operation location and land berth (Km).

We also use MH fleet concept here. That is, MHs are grouped in each collections of nD MHs and operates as a fleet. Each such group of MHs moves, travels and conducts (un) loading simultaneously. We assume that the number of MHs is a multiple of nD in this paper.

3. THROUGHPUT ANALYSIS

In this section, several techniques to calculate the MH system capacity and ship service time are introduced. Here, we assume that containership inter-arrival times are deterministic. When a ship arrives, all resources in the system are devoted to serve the ship. We obtain the ship service time using three techniques; there are differences between them. In section 3.1, assuming sufficient berth resources, we calculate a bound on the number of MH fleets required to provide ideal ship service. In section 3.2, we relax the assumption on sufficient berth resources. An approximation for service time term will be proposed. In section 3.3, our Petri-net model will be introduced. If we assume there are no containers to load, we can obtain the exact ship service time. Based on the Petri-net model, we can obtain the system throughput capacity.

3.1 Bound on the number of mobile harbor fleets: Ideal ship service

In this section, we obtain a lower bound on the number of mobile harbor fleets (MH fleets) required to provide uninterrupted service to incoming ships. There are several assumptions. First, loading begins after the unloading is complete. Second, all unloading and loading operations must finish before the next container ship arrives (We relax this assumption later.). Third, each group of nD MHs operates as fleet. Fourth, ship interarrival times are deterministic and there is never contention for a land berth. Fifth, number of containers to unload and load is equal (nTU); the ship enters and leaves with the same amount of containers. Sixth, the total loading and total unloading time are equal. When the first MH fleets is full, it leaves the container ship to release the containers at the port. As soon as the first MH fleet leaves, the next MH fleets starts unloading. The deterministic variables U, L, Tr, R, N_c^U, and N_c^L indicate unloading time ($\frac{cMH}{V_s}$), loading time ($\frac{cMH}{V_s}$), traveling time ($\frac{d}{V_t}$), releasing time ($\frac{cMH}{V_b}$), number of containers to unload, and number of containers to load, respectively. For simplicity, we assume that $N_c^U = N_c^L$ = nTU. When the total number of containers (both load and unload) is less than or equal to the MH fleet capacity, the number of mobile harbor fleets is equal to the container amount divided by the container capacity of a MH fleet, rounded up to the nearest integer. When the container amount is larger than a MH fleet capacity, if unloading time is longer than traveling time and releasing time (U≥2Tr+R), we need only two MH fleets, because they can alternate to provide uninterrupted services. When the unloading time is shorter than the traveling and releasing time (U<2Tr+R), we need more than two MH fleets. There are three cases for this condition. First is when the initial MH fleets returns before the entire unloading operations end. Second is when the initial MH fleet returns after all the unloading operations are complete, but before all the unloading and loading operations end. Last is when the initial MH fleet returns after all the unloading and loading operations are complete. The bound on the number of MHs required to provide continuous service is

$$
nMH \geq \begin{cases} 2, & \text{if } U \geq 2TR+R, \\ \min\left(\left\lceil \frac{nTU}{cMH \cdot nD} \right\rceil, \left\lceil \frac{2Tr+R}{U} \right\rceil + 1\right), & \text{if } U + 2Tr + R \leq \frac{nTU}{V_s \cdot nD} \\ \min\left(\left\lceil \frac{nTU}{cMH \cdot nD} \right\rceil, \left\lceil \frac{2Tr+R}{U} \right\rceil + 1\right) + \left\lceil \frac{(U+2Tr+R)-(\frac{nTU}{cMH \cdot nD} * U)}{L} \right\rceil, & \text{if } \frac{nTU}{V_s \cdot nD} \leq U+2Tr+R \leq \frac{nTU}{V_s \cdot nD} + \frac{nTU}{V_s \cdot nD}, \\ 2 * \min\left(\left\lceil \frac{nTU}{cMH \cdot nD} \right\rceil, \left\lceil \frac{2Tr+R}{U} \right\rceil + 1\right) + \left\lceil \frac{(U+2Tr+R)-(\frac{nTU}{cMH \cdot nD} * U + \frac{nTU}{cMH \cdot nD} * L)}{U} \right\rceil, & \text{if } \frac{nTU}{V_s \cdot nD} + \frac{nTU}{V_s \cdot nD} \leq U+2Tr+R, \end{cases}
$$

(1)

We will explain the third case in (1) to help understand the bound. Under the condition that the initial MH fleet returns after the unloading is complete but before the entire unloading and loading operations end ($\frac{nTU}{V_s \cdot nD} \leq$ U+2Tr+R $\leq \frac{nTU}{V_s \cdot nD} + \frac{nT_L}{V_s \cdot nD}$). Let's assume that cMH = 200TEU, nTU = 5000 TEU, nD = 6, V_s = 50 TEU/hr, Tr = 9, and R = 4. Then, $\frac{nTU}{V_s \cdot nD}$ = 16.67 ≤ U+2TR+R = 26 ≤ $\frac{nTU}{V_s \cdot nD} + \frac{nT_L}{V_s \cdot nD}$ = 33.34. This means that the berth is located farther, so that the MH fleet returns to the ship after the unloading, but before the loading operations are complete. If we insert the numbers to the equation, $\left\lceil \frac{nTU}{cMH \cdot nD} \right\rceil$ = 5, $\left\lceil \frac{2Tr+R}{U} \right\rceil + 1$ = 7, and minimum of these is 5. So we need 5 MH fleets to unload and an extra 2 MH fleets for loading ($\left\lceil \frac{(U+2Tr+R)-(\frac{nTU}{cMH \cdot nD} * U)}{L} \right\rceil$), because as soon as the last unloading MH fleet leaves to the berth, we need alternate MH fleet for loading. This is obtained by subtracting total time to unload ($\frac{nTU}{cMH \cdot nD} * U$) from time a fleet takes to come back to the ship (U+2Tr+R) and divide by loading time (L). Therefore, at least 7 MH fleets are needed to complete the service.

In addition to these equations, we consider the situation where another ship arrives before the service is complete for the previous ship. There are two situations. One is when the next ship arrives before the entire unloading operations of the previous ship are complete and the other is when the next ship arrives after the entire unloading operations are done but before all the loading operations of the previous ship are complete. For the first situation, additional MH fleets are needed to conduct service for the next ship. To calculate the multiple of MH fleets, we need to divide total unloading time by inter-arrival time. In case of second situation, only one additional of MH fleets is required, because the MH fleets used in the previous ship to load the containers to ship can be used again.

Using these relationships, we can obtain an approximation for the service time as (nMH x $\frac{2cMH \cdot nD}{V_s}$). In this section, the number of mobile harbors is calculated with an assumption that all the unloading and loading operations are completed within the inter-arrival time and therefore, it is certain that there will be no waiting time for the ships. This method clearly shows the relationship between size of the mobile harbor fleet, distance between the ship and berth, and the number of mobile harbor fleets. If the size of the mobile harbor fleets and distance between the ship and berth are given, the lower bound of the number of MH fleets can be easily obtained. However, there exist several restrictions. First, there may not be sufficient berths to guarantee no MH contention at the land berths. Second, mooring time is not considered. In the next section, a method to calculate the number of MHs considering the restrictions stated above will be proposed.

3.2 Ship service time approximation: MH delays at the land berth and ships

Relaxing the assumptions of section 3.1, we develop a ship service time (SST) approximation with finite berth resources. There are three container flows in the mobile harbor system. First, containers are transferred from the containership to the MH at sea via the MH container crane. Second, containers travel from the offshore operation location to the land berth on the MHs. Third, containers are transferred from the MH to the land at the MH-land berth interface. The maximum container throughput is less than the minimum capacity at each of three flows. That is the system capacity is less than (or equal) to the system bottleneck capacity. Let λ_1, λ_2, λ_3 denote the cargo flow rate between ship and offshore operation location, cargo flow rate between offshore operation location and MH berth and cargo flow rate between MH berth and land berth.

$$\mu \leq \min\{\lambda_1, \lambda_2, \lambda_3\} =: \mu^*$$

A first order approximation for the SST is

$$\text{SST} \approx \frac{nTU}{\mu^*} \qquad (2)$$

However, MHs may suffer queuing at the land berths and ships if their number is greater than the number of docking points or berths. The queuing also depends on the arrival pattern of MHs to the berths and docking points. As a gross approximation to this queuing, we assume an arrival pattern of MHs to the berths and docking points as in figure 2.

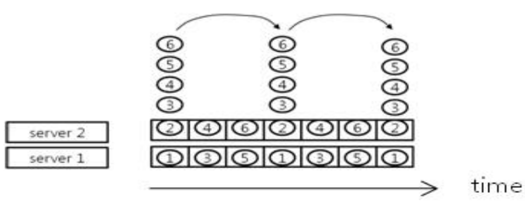


Figure 2. The assumed arrived pattern of MHs to their servers

All MHs come to the server (which can be berth or docking point) simultaneously. The server immediate begins service of as many MHs as possible. In figure 2, MHs 1 and 2 enter service with the two servers immediately. The remaining MHs in the fleet 3,4,5,6 in figure 2 are queued until MHs 1 and 2 complete services. MHs 3,4 and 5,6 enter service immediately as they are able. When MHs 5 and 6 have received service the process repeats with a new MH fleet. In this case, the average waiting time for MHs to enter service is

$$T_w = \frac{N_S * T * \sum_1^{\left\lfloor \frac{Nc}{Ns} \right\rfloor - 1} n + \left\{ \left(Nc - \left\lfloor \frac{Nc}{Ns} \right\rfloor * Ns \right) * \left\lfloor \frac{Nc}{Ns} \right\rfloor * T \right\}}{Nc} \quad (3)$$

where, Nc is the number of MHs in a fleet, Ns is the number of servers and T is the service time. Thus, the capacity of the second flow between the ocean location and the land berth location, becomes $\frac{cTU}{\frac{cMH}{Vs} + 2 * \frac{d}{Vt} + \frac{cMH}{Vb} + Twb + Tws} * nMH$, in which Twb and Tws are the average waiting time of MHs at the berth and ship respectively. Using (3) and the approximate SST becomes

$$SST \approx \frac{nTU}{\min\left(\min(nD, nMH) * Vs, \frac{cTU}{\frac{cMH}{Vs} + 2 * \frac{d}{Vt} + \frac{cMH}{Vb} + Twb + Tws} * nMH, \min(nB, nMH) * Vb\right)} \quad (4)$$

The result of this equation was compared with exact simulation under the same assumptions. We compared 154 systems and the average absolute gap between the SST approximation and the simulation was 6.49%. The approximation is intuitive and generates reasonable results. Also, one can easily obtain the service time distribution of the system from the distribution of the cargo volume, since the denominator of (4) is constant given the system parameters.

3.3 Petri net model

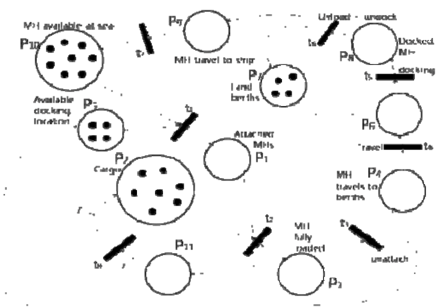


Figure 3. Petri-net model

Petri net is a graphical and mathematical modeling tool that can be used for many systems, see Tadao Murata (1989). In C.V. Ramamoorthy and Gary S. Ho(1980) deterministic timed nets and their minimum cycle time are discussed. A Petri net consists of circles that represent conditions and bars that represent events. The circles are called places and the bars are called transitions. If each input place of a transition has at least one token (dot), the transition fires and a token is removed from each of the input places and a token appears in each of the output places. We assume that transitions can fire again before the previous firing is complete, so long as there are sufficient tokens available. That is, the firings of a transition can overlap in time. Here we develop a Petri net model of a port with mobile harbors. The model is depicted in Figure 3; we describe its transitions and places next. A mobile harbor attaches to a ship docking point during transition t_1. Then, it loads cargo during transition t_2 and detaches from the ship during t_3. A token indicating docking location returns to its place when the mobile harbor is detached. Travel to the land berth occurs during transition t_4. It docks to a land berth during t_5 and unloads the cargo. After undocking (t_6), the token arriving at place P_7 represents a land berth being ready for another mobile harbor full of cargo. The unloaded and undocked mobile harbor travels back to the ship and repeats the cycle. Transitions t_2 and t_8 model the cyclic arrival of subsequent ships – that is, another ship enters the service after all cargo has been removed from the previous ship. In Fig. 3, the tokens indicate that there are 8 mobile harbors, 7 full mobile harbors worth of cargo on each ship, 4 berths, and 4 docking points on a ship. Note that t_8 in Fig. 3 requires 7 tokens to fire. As a consequence, when every single token arc fires 7 times, t_8 fires once.

If the system is decision-free, the cycle time can be readily obtained analytically. A decision-free Petri net is a net in which each place has only one input arc and one output arc. Our model in Figure 3 is decision-free. As described in C.V. Ramamoorthy and Gary S. Ho(1980), the cycle time C for a decision-free Petri net (assuming all transitions fire immediately as they are enabled) can be obtained as

$$C = \max\left\{\frac{T_k}{N_k} : k = 1, 2, \cdots, q\right\},$$

where k is the index of circuits (all directed paths from each place back to itself), T_k, the sum of the execution times of the transitions in circuit k, is $\sum_{t_i \in L_k} r_i$, and r_i is the firing duration of transition t_i. N_k, the total number of tokens in the places in circuit k, is $\sum_{P_i \in L_k} M_i$, where M_i is the number of tokens in place P_i. Thus, the cycle time C is

$$\max\left\{\begin{array}{c}\frac{r_1+r_2+r_3}{nD}, \frac{r_2+r_8}{nTU}, \frac{r_1+r_2+r_3+r_4+r_5+r_6+r_7}{nMH}, \frac{r_5+r_6}{nB}, \frac{r_1+2r_2+r_3+r_4+r_5+r_6+r_7+r_8}{nMH+nTU}, \\ \frac{r_3+r_1+2r_2+r_3+r_8}{nTU+nD}, \frac{2r_1+2r_2+2r_3+r_4+r_5+r_6+r_7}{nD+nMH}, \frac{r_1+r_2+r_3+r_4+2r_5+2r_6+r_7}{nMH+nB}, \\ \frac{2r_1+3r_2+2r_3+r_4+r_5+r_6+r_7+r_8}{nD+nMH+nTU}, \frac{2r_1+2r_2+2r_3+r_4+2r_5+2r_6+r_7}{nB+nMH+nD}, \\ \frac{r_1+2r_2+r_3+r_4+2r_5+2r_6+r_7+r_8}{nB+nMH+nTU}, \frac{2r_1+3r_2+2r_3+r_4+2r_5+2r_6+r_7+r_8}{nD+nMH+nB+nTU}\end{array}\right\} \quad (5)$$

Cycle time can be interpreted as the time between executions for each transition. The throughput of the system is the inverse of the minimum cycle time (1/C). That is, one full mobile harbor load of cargo can be moved to land per minimum cycle time C. Since another ship arrives to the system when all cargos are removed from the previous ship, the total service time of a ship is the cycle time times the number of cargo tokens. The petri net analysis provides an exact value for the throughput. However, it can be difficult to model the complexity of a MH system as a decision free Petri net.

4. QUEUEING ANALYSIS

In the previous section, we developed several techniques to calculate system throughput and ship service time. There we assumed deterministic ship arrival and other ideal conditions. However, under practical condition a ship may wait until resources are available. Here, we conduct queuing analysis to assess such delay. In section 4.1, we employ a traditional M/M/c model to estimate the service level. Using the service level information, we can determine an appropriate number of MHs and berths. In section 4.2, an approximation for the cycle time of a ship is proposed. To verify the approximation, we compare it with simulation.

4.1 Approximation for the number of MHs and berths to achieve a given service level

If we assume that the distribution of the ship interarrival times (IAT) and cargo volumes possess an exponential distribution, an M/M/c queuing model can be used. There, the service time distribution of a ship is approximated as exponentially distributed based on our SST approximation of (4) (since the denominator is constant and the numerator we have assumed exponentially distributed). Also, we assume MHs and berths form an exclusive group to service a ship and MHs always work with berths in the same group. A group is dedicated to a ship and another group may serve that ship. The port can employ several such MH and berth groups. Since explicit analytic solutions for the M/M/c queue, we can readily calculate the mean queuing time an arriving ship will experience in this model. This in turn allows us to obtain a commonly employed service level metric: the waiting time to service time ratio (W/S). We illustrate the application of this approach via example.

Example 1: Assume that containership arrives to a port deterministically with a 12 hour interarrival time. Let nTU= 2500 TEU, time to travel is 30 min, ocean operation speed is 45TEU/hour and land operation speed is 120TEU/hour. There are three groups of MHs and berths. The W/S ratio based on the resulting M/M/c queuing model is shown in Table 1. There we vary the number of berths and MH units in each of the three groups (each group is the same size). The shadowed region shows the number of berth/group and MHs/group combinations that provide W/S ratio exceeding 10%. The entries in bold show the frontier at which W/S ratio is less than 10%; such values indicate the promising resource combinations to achieve good service. They should next be compared in terms of cost. ■

Table 1. W/S ratio of example market

		Number of berth/group						
		1	2	3	4	5	6	7
Number of TU/group	8	0.458552	0.187933	0.140075	0.116523	0.109413	0.102645	**0.096204**
	9	0.420892	0.173922	0.124978	0.107745	**0.097327**	0.092416	0.087696
	10	0.380961	0.154699	0.113226	**0.096094**	0.084603	0.081013	0.077537

	11	0.344007	0.139623	0.1004	0.084318	0.075648	0.070301	0.070301
	12	0.311507	0.124108	**0.08819**	0.073517	0.070301	0.070301	0.070301
	13	0.299763	0.11957	0.084887	0.070712	0.070301	0.070301	0.070301
	14	0.285772	0.112604	0.079958	0.070301	0.070301	0.070301	0.070301
	15	0.271129	0.106508	0.074328	0.070301	0.070301	0.070301	0.070301
	16	0.259436	**0.099532**	0.070301	0.070301	0.070301	0.070301	0.070301

Our method can be used to obtain an initial design for a potential market. We can propose a promising number of resources for a given market condition with a required W/S ratio level. However, the grouping assumption may overestimate the required number of resources since it has less flexibility than a real system.

4.2 Approximation for cycle time of a ship

Figure 4. Modified M/G/c queuing system

In this section, we propose an approximation formula for the cycle time of a ship in a mobile harbor system. Cycle time includes waiting time and service time of a ship. We again employ the fleet concept and group the MHs into collections containing nD MHs. Here we also assume that berths act in groups of size nD. There will practically be fewer such berth groups than mobile harbor fleets. When a ship arrives to the offshore operation location, a mobile harbor fleet moves to the ship and conducts the unloading operation. After the MH fleet is filled with containers, the fleet moves to a land berth group and releases containers simultaneously to the land berths. When there are no containers remaining to transfer from the ship to the berth, the ship exits the system.

For analytical simplicity, we assume that once a ship begins service with a MH fleet, it will received dedicated service from that fleet and no other, even if the ship is waiting for the fleet to return from a trip to the land berth and another MH fleet is available. This simplification will cause error in low loading, but as the system becomes busier, there will be few other available MH fleets. There are two key situations in which resource contention waiting can occur. First, an arriving ship has to wait when there are no available MH fleets. Second, MH fleets have to wait when all land berths groups are full. Considering each MH fleet as a server and assuming that cargo ships arrive as a Poisson process, we can model the ship queue as an M/G/c system whose service time is the cycle time of an M/D/m queue (representing the time for a MH fleet to take containers to the land berth including queuing at the berths). We also assume that the number of containers in a ship has exponential distribution with mean nTU. All other variables are as in section 2. Figure 4 depicts the system; there, the number of docking points nD is equal to 2. In figure 4, two mobile harbors act as one server (a fleet). When a ship arrives to the system, if there are available servers, the ship occupies one server and its two mobile harbors conduct unloading operations simultaneously. After the MH fleet is filled with containers, the MH fleet moves to a land berth group and releases containers to the land berths. It subsequently returns to the offshore operation location. If there are no containers remaining to transfer, the ship exits the system.

Employing an approximation introduced in Morrison and Martin(2007), we develop an approximation for the cycle time of the system. First, we estimate the ship service time. Treating the MH visits to the land berth as an M/D/m queue (m is the number of berth groups), the ship service time approximation is:

$$E[SST] \approx \frac{nTU}{nD\times Vs} + \left\lceil \frac{nTU}{nD\times cMH} - 1 \right\rceil \times \frac{2d}{Vt} + \frac{nTU}{Vb} + \frac{1}{2}\frac{nC}{Vb}\frac{{\rho_b}^m}{(1-{\rho_b}^m)} \ , \tag{6}$$

where the berth loading ρ_b is $(nC/Vb)(\lambda/m)$, and m is the number of berth groups (nB/nD). The symbol $\lceil p \rceil$ is the smallest integer greater than or equal to p. The first term represents the ship unloading time at the offshore operation location. The second term represents MH fleet travel time to the berth. The third term represents unloading time at the land berth. The last term represents queuing time in the case that there is no available land berth when the MH fleet arrives. Employing this service time, we propose an approximation for the ship cycle time as follows:

$$E[CT] \approx E[SST] + \left(E[SST] + \frac{2d}{v_t}\right)\left(\frac{1+C_S^2}{2}\right)\left(\frac{\rho_s{}^n}{(1-\rho_s{}^n)}\right), \tag{7}$$

where n is the number of MH fleets (nMH/nD) and $C_S{}^2$, the squared coefficient of variation, is approximately $\frac{\lambda^2\rho_b(2-\rho_b)}{(\frac{Vb}{nC}-\lambda)^2}$ (since the SST has been approximated by an M/D/m queue, this represents the coefficient of variation (COV) of the cycle time of an M/D/m queue. For simplicity, we approximate this by the COV of an M/M/1 but correct the loading to account for the multi servers.). The system loading ρ_S is $\lambda(1/n)*(E[SST]+(2d/v_t))$. To verify the approximation, we conduct simulation (nD=4, nTE~Exp(1/1565) and λ=(1/8.6)). The result is below:

Table 3. Performance of the approximation

Group of berths	# of MH fleets	Simulation	Approximation	Error
4	8	60.60	64.30	6.11%
8	8	58.77	58.62	0.26%
4	12	23.72	23.10	2.61%
8	12	23.63	22.58	4.44%
4	16	17.31	16.99	1.85%
8	16	16.86	16.15	4.21%
4	20	14.53	15.23	4.81%
8	20	14.12	14.73	4.32%

The errors are in 10%. We can see that approximation is sufficient.

5. CONCLUSION

In this paper, we develop several techniques to evaluate the mobile harbor system. Based on throughput analysis, we obtain a lower bound on the number of MHs required to provide ideal ship service. An approximation for ship service time which relaxes assumptions of the previous approach is also introduced. We develop a Petri-net model to calculate throughput and cycle time values when there is only unloading. Then we conduct queueing analysis to estimate W/S ratio and show how this can be used for the initial design in a potential market. We also develop an approximation for the cycle time of a ship. Using a modified M/G/c queue, we develop an approximation for the queuing and service time of container ships. The errors of approximations are smaller than 10%. We also develop simulations of port systems with offshore components to test the approximation terms. These analysis techniques may be improved by consideration of how they may support each other; it is our future direction.

6. REFERENCES

Drewry Shipping Consultants. (2006). *The Drewry Annual Container Market Review and Forecast 2006/2007*. Drewry Shipping Consultants, London.

N. P. Suh. (2008). *Mobile Harbor to Improve Ocean Transportation System*. Korean Patent Application No. 2008-12428.

N. P. Suh. (2008). *Mobile Harbor to Improve Ocean Transportation System*. Korean Patent Application No. 2008-41981.

Isobe, E. (1999). Research and Development of Mega-float. *Proceedings of the International workshop on Very Large Floating Structures* (VLFS' 99), 7-13.

Tadao Murata. (1989). Petri Nets: Properties, Analysis and Applications. *Proceedings of the IEEE*, 77: 541-580.

C.V. Ramamoorthy and Gary S, Ho. (1980). Performance Evaluation of Asynchronous Concurrent Systems Using Petri Nets. *IEEE Transactions on Software Engineering, SE-6*: 440-449.

Jonghoe Kim. (2010). *Generation & Evaluation of Mobile Harbor Architectures.* M.S thesis, Korea Advanced Institute of Science and Technology.

James R. Morrison and Donald P. Martin. (2007). Practical Extensions to Cycle-time Approximations for the G/G/m Queue with Applications. *IEEE Transactions on Automation Science and Engineering, 4:* 523-532.

THE SIMULATION MODELS AND TECHNOLOGIES FOR PORT LOGISTIECS BASED ON EMERGENCY MANAGEMENT

Yan-Ping Meng1*, You-Fang Huang1*, Zhi-Hua Hu2*, Bing Yang2*, and Yi Ding2*

Logistics Research Center
Shanghai Maritime University
ShangHai, 201306, China
E-mail: ypmeng2008@gmail.com

Abstract: As ports have become key logistics nodes in global chains, and Port logistics plays important role in world trade. The emergency management of port logistics is critical and important with the frequent threatens of natural and human disasters. A number of modeling and simulation tools have been developed and more are being developed for port logistics. The simulation models and technologies for Port Logistics were analyzed in this work, and the ones which are suitable for emergency management of port logistics were summarized from different aspects of application. The types and characteristics emergency management of port logistics are concluded and the corresponding simulation models and technologies were suggested.

1. INTRODUCTION

The development of global supply chains coupled with globalization of production and consumption and increased demand for raw materials and finished goods have significant impacts on the role of ports. Ports have become critical parts of global supply chains A wide variety of industries rely on efficient port logistics operations, the efficiency port logistics is important to international trade and foreign direct investment (FDI), and thus to economic growth, so both the governments and goods owners always pay a great attention to port logistics.

As such, when a port cannot operate because of a crisis, such as a natural disaster (e.g. earthquakes, hurricanes, flooding) or man-made crisis (e.g. terrorist attacks, fires) (Tugce G. Martagan, 2009), it will has negative impact on the port logistics and influence a series of participators.

In view of the new security regulations and the requirements of increased visibility throughout the supply chain, emergency management is becoming a significant operational challenge that international logistics and supply chain operators face today. The emergency management of port logistics have gradually attracted attention from both industry and academia domains. An International Workshop on Risk Management in Port Operations, Logistics and Supply Chain Security was hold in September 2006 to share the experience and push the further and deeper research.

Emergency management is a complex policy subsystem that involves an intergovernmental, multi-phased effort to mitigate, prepare for, respond to, and recover from disasters (William J. Petak, 1985). Emergency management of port logistics is complicated and dynamic because of highly uncertain nature of risk. As simulation refers to a model built in order to determine the response of a system to changes in its internal structure and inputs (C.R. Harrell, 1996), and usually it consists of a simplified representation of a dynamic process deemed too complex for direct analysis, so simulation provides a good tool to analysis and modeling the crisis and decision making support for emergency management of port logistics.

2. SIMULATION MODELS AND TECHNIQUES

2.1. Simulation models and techniques

A simulation is the imitation of the operation of a real-world process or system over time; it is a popular technique of modeling (A.M. Law, 2000).Many real-world systems are so complex that models of these systems are virtually impossible to be solved mathematically. In these instances, numerical, computer-based simulation can be used to imitate the system. Usually, a simulation models consists of a simplified representation of a dynamic process deemed too complex for direct analysis. The simulation model is developed to describe the behavior of a system in its internal structure and inputs and usually takes a set of assumptions which are expressed in mathematical, logical and symbolic relationships between entities, or objects of interest, of the system. Thus, the simulation model is smaller, cheaper and simpler than doing the experiments on the real system, and its purpose is to analyze and understand the behavior of a system as a function of actions and alternative scenarios (Jerry Banks, 2009) .

Simulation models can be used both as an analysis tool for predicting the effect of changes to existing system with a wide variety of "what-if" questions about the real-world system, and as a design tool to predict the performance of new system in the design stage under varying sets of circumstances . Simulation methodologies are not specific for a certain application, but, rather, have general characteristics that allow their use, after adaptation, to almost any problem. However, they are especially useful in situations in which the intrinsic stochastic characteristics of a problem, along with the interaction of its elements, make the use of classical statistical techniques unsuitable.

2.2. Simulation Application

Simulation supports the analysis, proposal and optimization of real systems by: Substitution of the real system by the simulation model; Experimentation on the simulation model through discovering its properties, behavior and reaction to different conditions; Application of the obtained results to the real system both existing or planned (Kavička, A.2005).

The application of simulation is vast, such as manufacturing, business processing, constructing engineering and project management, logistics, transportation and distribution and so on. Recent studies show that simulation has already used a lot in port logistics, however, the majority of them are dealing with the normal conditions. The following issues belong to the most typical problems of port logistics being solved utilizing simulation techniques: (a)Rationalization of operation processes, such as the simulation study of sequencing delivery and receiving operations for yard cranes to compare the performances of the suggested approaches (Kap Hwan Kim et al. (2003)), the logistic chain simulation in the northwestern Italian port system for highlighting both features and problems of the intermodal network(Francesco Parola et al. (2005)) ; (b)The cost savings or capacity increasing through optimum use of resources, such as the simulation model of the terminal work with strategy to increase the capacity of terminal for bulk cargo unloading (U. Bugaric, et al. (2007)); (c)Infrastructure planning, Reconstructing and configuring verification, such as analysis and comparing all hub scenarios in the handling of barges at potential hubs and at sea terminals on waiting times and capacity utilization by simulation (A. Caris et al. (2010)); (d)Improvement of operation control strategies, such as the simulation model of the port of Thunder Bay to assess future management strategies (Said M. Easa (1986)); (e)Crisis situations management or disruption recovery, such as the simulation for interruption of cargo handling countermeasure for ship mooring problems (Shigeki Sakakibara, et al. (2008)). Simulation with crisis conditions is attracting more and more attention and the level of disaster preparedness is beginning to be improved.

3. EMERGENCY MANAGEMENT IN PORT LOGISTICS

3.1. Emergency management becomes hot

The modern world is a high risk society. Since the profound impact of 9/11 attacks on the United States, the profession and practice of emergency management has attracted a lot of attention. Emergency management and prepare for and respond to catastrophic disasters. Except the threat of terrorism, there are more certain risk posed by hurricanes, earthquakes, wildfire, and other familiar hazards. Ports are the primary connection between the countries and the world economy. Maritime safety and Emergency management in Port logistics begin to significant concerning of lots of the agencies. For example, the U.S. Coast Guard has strengthened the maritime safety, the Maritime Transportation Security Act (MTSA) focuses to prevent security disruptions in the maritime supply chain, i.e., in the movement of maritime cargo from shipper to consignee, with the emphasis on the port link of the chain (C. Ariel Pinto, Wayne K. Talley, 2005). The research community has also responded to the events of disaster both form nature and man-made crisis in port logistics. Some of recent studies have investigated risk and security in port logistics and maritime supply chain.

3.2. Emergency management in port logistics

Port logistics are tightly connected networks all around the world and the global impacts of threats intensify with the influence expanding, from natural disasters, terrorist attacks, wars, worker strikes, etc. The strategies of emergency management for preparation, responding and recovery depend highly on the specific characteristics of the types of the crisis, and are designed based on experience and analysis of event records.

C. Ariel Pinto and Wayne K. Talley (2005) divided the crisis in port logistics into two types, the accidental disruptions and potential security disruptions. They did an investigation of port logistics accidents in 95 countries to classify the port logistics related accidents by type, origin and cause and location. The origins of the accidents can be separate into 3 groups: transport of cargo, loading and unloading operations, other origins (storage, waste and warehouse facilities). The types the accidents in port logistics include the releases and loss of containment, fires, and explosions. The location of accidents occurrence be divided into 3 kinds: the sea during approach and maneuver; on land during storage, process and transport; and at a sea-land interface during loading/unloading and maintenance. The potential security disruptions, in particularly, means the possible terrorism related disruptions such as nuclear attacks, poison gas attacks, dirty bomb explosions, or commandeered container trucks and ships have not been recorded yet.

The most common types of substances in port accidents are crude oil and other hazardous products. Port logistics accidents that involve the handling and temporary storage of hazardous cargo in port areas originate from the port activities, such as hardware failures of ship, inland and loading/unloading equipment and external events (e.g. bad weather).

4. SIMULATION FOR PORT LOGISTICS BASED OF EMERGENCY MANAGEMENT

4.1. Simulation for emergency management

Simulation for emergency management is beginning to be the hot topic with the development of simulation technology and the huge impact and importance of emergency management. The winter simulation conference (WSC) is the main publisher of the newest simulation technology and applications. WSC had topics of Homeland Security/ Emergency Response and risk analysis in these years. For example, Bo Zhang et al. (2009) presented an agent-based model for hurricane evacuation which captures the interactions among evacuees during the evacuation process, and simulate different scenarios on a test network. Md Yusuf Sarwar Uddin et al. (2009) did research on the post-disaster mobility model for delay tolerant networking. Young M. Lee et al. (2009) simulated the distribution of Emergency Supplies for Disaster Response Operations.

4.2. Simulation areas for port logistics based of emergency management

The Simulation for port logistics based of emergency management is still on it infancy stage, but the safe and security problems are attacking more and more concerning. A. G. Bruzzone et al. (2000) proposed the development of an integrated interactive environment for risk analysis applied to harbor and maritime environments which is very effective for properly designing the harbor and maritime infrastructure in terms of resources, structures and services for facing possible emergencies. Harrald et al. (2004) proposed a framework for sustainable port security, including port-security prevention (pre-attack) and mitigation and recovery (consequence) programs.

There are some simulation models focusing on special disaster or crisis in port logistics. crude oil spilling and hazardous article storage and transportation are the most frequent research objects. Massimo Paolucci et al. (2002) applied a simulation-based approach to solve the problem of allocating the crude oil loads of tanker ships to port and refinery tanks (PRT). E. Palazzi et al. (2004) built models for risk assessment of hydrocarbon spills in port area. Malcolm L. Spaulding et al. (2007) developed a fully integrated, geographic information based modular system simulations for a tanker entering Block Island Sound to illustrate the model predictive capability for realistic emergency scenarios.

Earthquake, flooding and cliff falls are also the common types of disasters in port logistics. K. F. Cheung, et al. (2003) described a model package that simulates coastal flooding resulting from storm surge and waves, which applied to the coastal flooding caused by Hurricanes Iwa and Iniki that hit the Hawaiian Island of Kauai in 1982 and 1992, respectively. Byung Ho Choi, et al. (2008) proposed a three-dimensional simulation of the 1983 central East (Japan) Sea earthquake tsunami at the Imwon Port (Korea). Jim W. Hall, et al. (2002) introduced a episodic stochastic simulation model to model the duration between cliff falls as a gamma process and fall size as a log-normal distribution.

There are some special crisis in particular regions and ports, for example, the accelerated growth of brash ice is a problem that the port operator has to deal with in busy cold region harbour basins. Huachen Pan, et al. (2009) used flow and heat transfer simulations to investigate the ice control design options for the new Helsinki Vuosaari harbour, and studied five design options with different warm water supply locations and bubbler arrangements using numerical simulations.

From the papers published on the simulation for port logistics based on emergency management, we can found the simulation areas for port logistics based on emergency management includes 3 types: disaster or crisis engineering simulation; the crisis analysis or assessment simulation based on different scenarios, and the emergency response or recovery strategies simulation. And the simulation will offer a great tool to help emergency decision supporting tool for the port society.

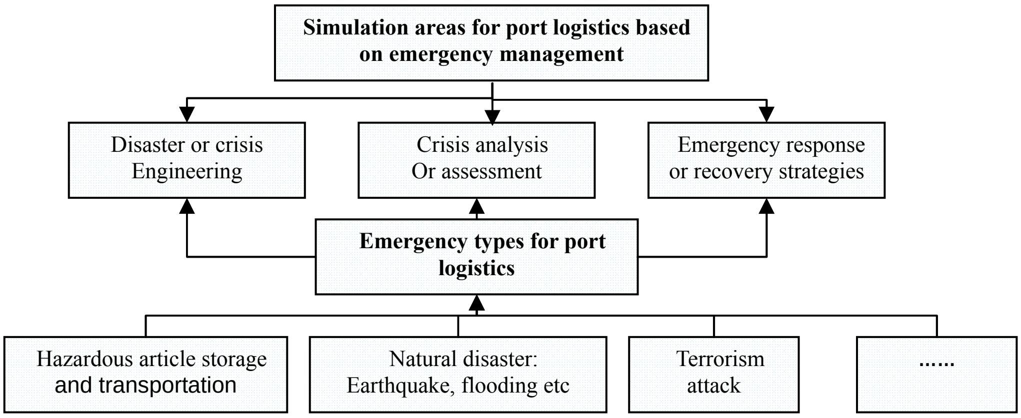


Figure 1 Simulation areas for port logistics based on emergency management

4.3. Simulation approaches for port logistics based of emergency management

Some approaches of simulation used in general emergency management can also be applied in emergency management of port logistics, such as simulation-optimisation, distributed simulation, and so on. Rui C. Guimarães et al. (1989) presented Simulation-optimization as a method for solving complex decision problems. Josefa Z. Hernández et al. (2001) developed Knowledge-based models for emergency management systems. Wan Hu et al. (2008) proposed a distributed computer simulation framework which realizes extending HLA/RTI to Internet based on Grid service for helping to perfect the disaster response plan.

The spatial technology and visual technology, such as GIS, visual reality, are integrated into the simulation models in emergency management to offer user-friendly graphical interface. Feng Xu et al. (2008) applied GIS, CAD, FEA, and VR integrated to earthquake disaster Simulation for an Urban Area. M. Kulawiak et al. (2010) implemented a Web-based Geographic Information System (Web GIS) for an existing oil spill monitoring and forecasting service, allowing authenticated end users to view the simulation results in a geographical context. Ameya Shendarkar et al. (2008) presented a VR-based Belief, Desire, and Intention (BDI) software agent to construct crowd simulation under terrorist bomb attacks in public areas which can be used as an effective emergency management tool.

Songbai Cheng et al.(2009) advanced a dynamic simulation model of toxic gas diffusion, and developed a 3D gas diffusion visualization scene management subsystem including leak hazard source (tank), leak background (ground and sky) and surroundings near the leak area, based on virtual reality (VR) technology, providing references for accident impact prediction, assessment and emergency plan. Albert Douma et al. (2009) proposed a multi-agent based approach for enabling distributed planning of container barge rotations along terminals in the port of Rotterdam, and examined the value of exchanging different levels of information and evaluate the performance by means of simulation. W. Hall, et al. (2002) introduced a stochastic simulation model to model the cliff falls, which is often preferable to conventional regression models.

4.4. Matters need attention and the trends

There are several matters which need attention during the modeling process for port logistics based on emergency management. First, the availability of related data for simulation is important. As most simulation models, the validity of the output will be heavily depended on the quality of the input variables. Most simulation for port logistics based on emergency management is grounded on the crisis record files combining the data collecting from the operation. As the increased public awareness of emergency management in port logistics, data collection for simulation is beginning more and more optimistic. Second, the simulation model must focus on the purpose, and capture the main characteristics and influence factors of the object system. Some research result proved that relatively simple simulation modeling can provide valuable support to the decision making process (Alan C. Cowdale, 2003). To achieve this effect, the simulation staffs and experts should be directly involved in the discussions surrounding the key decisions. Third, the simulation models should analysis multi-scenarios of the crisis and assess different recovery strategies, and could be modified to meet the developing requirements during the disaster preparation, responding, and recovery.

From the papers reviewed, the trends of simulation for port logistics based on emergency management can be concluded as flows: (a) Graphical modeling, visual technology provides the end user with comprehensive and synthetic, both spatial and temporal, environmental information through a remotely customizable user-friendly graphical interface; (b) Interactive and net communication: emergency management of port logistics involves a lot of organization, the communication and information sharing is important, so the simulation should support this future; (c)

Objective-oriented modeling will bring convenience for simulation modifying according to the dynamic development of the situation; (d) Distributed agents technology, as the emergency problem is complex and extensive, distributed agents will strengthen the capability to dealing with the problem and accelerate the simulation speed.

5. SUMMARY AND CONCLUSION

Port logistics plays important role in world trade. With the frequent threatens of natural and human disasters, the emergency management of port logistics is critical and important. In this paper, the simulation models and technologies for port logistics were analyzed, and the main areas and approaches for port logistics based on emergency management were summarized. According to the types and characteristics of emergency management in port logistics, some suggestions for attention in simulation were proposed, and some trends of emergency management simulation were concluded.

ACKNOWLEDGMENT

This work was supported in part by Shanghai Nature Science Foundation Project (No. 10ZR1413200), Innovation Program of Shanghai Municipal Education Commission (No. 10YZ115), Shanghai Science Commission International Cooperation Project (No. 09530708200), Shanghai Science Commission Project (No. 09DZ2250400), Shanghai University Key Subject Construction Grant (No. J50604) and Shanghai Maritime University Graduate Student Innovation Grant (No.yc2009105).

5. REFERENCES

[1] Tugce G. Martagan, Burak Eksioglu, Sandra D. Eksioglu, Allen G. Greenwood, A Simulation Model of Port Operations During Crisis Conditions, *Proceedings of the 2009 Winter Simulation Conference*, 2832-2843

[2] William J. Petak, Emergency Management: *A Challenge for Public Administrator*[book], 1985

[3] C.R. Harrell, R.E. Bateman, T.J. Gogg, System improvement using simulation, Orem, Utah: PROMODEL Corporation, 1996.

[4] Jerry Banks, Barry L. Nelson, David M. Nicol, *Discrete-Event System Simulation*, 2009, books.google.com

[5] Kavička, A.; Klima, V. 2005. ABAsim: Agent-based architecture of simulation models, in Simulation Almanac, Czech and Slovak Simulation Society, Prague, 63-72.

[6] Kap Hwan Kim, Keung Mo Lee, Hark Hwang, Sequencing delivery and receiving operations for yard cranes in port container terminals, *International Journal of Production Economics*, 2003, 84(3): 283-292

[7] Francesco Parola, Anna Sciomachen, Intermodal container flows in a port system network: Analysis of possible growths via simulation models, *International Journal of Production Economics*, 2005, 97(1):75-88

[8] U. Bugaric, D. Petrovic, Increasing the capacity of terminal for bulk cargo unloading, *Simulation Modelling Practice and Theory*, 2007, 15(10): 1366-1381

[9] A. Caris, C. Macharis, G.K. Janssens, Network analysis of container barge transport in the port of Antwerp by means of simulation, *Journal of Transport Geography*, In Press, Corrected Proof, Available online 25 January 2010

[10] Said M. Easa, Assessing future management strategies in the port of Thunder Bay, *Transportation Research Part A*: 1986, 20(3): 185-195

[11] Shigeki Sakakibara, Masayoshi Kubo, Characteristics of low-frequency motions of ships moored inside ports and harbors on the basis of field observations, *Marine Structures*, 2008, 21(2-3): 196-223

[12] C. Ariel Pinto, Wayne K. Talley, Modeling of Commercial Maritime Port Recoverability from Security Disruption: Work-in-Progress, *26th ASEM National Conference Proceedings*, 2005, 470-474

[13] Bo Zhang, Satish Ukkusuri, Wai Kin (Victor) Chan, Agent-Based Modeling for Household Level Hurricane Evacuation, *Proceedings of the 2009 Winter Simulation Conference*, 2009, 2778-2784

[14] Md Yusuf Sarwar Uddin, David Nicol, Tarek Abdelzaher, Robin Kravets, A Post-Disaster Mobility Model for Delay Tolerant Networking, *Proceedings of the 2009 Winter Simulation Conference*, 2009, 2785-2796

[15] Young M. Lee, Soumyadip Ghosh, and Ettl Markus,Simulating Distribution of Emergency Supplies for Disaster Response Operations, *Proceedings of the 2009 Winter Simulation Conference*, 2009, 2797-2808

[16] A. G. Bruzzone, R. Mosca, R. Revetria, S. Rapallo, Risk analysis in harbor environments using simulation, *Safety Science*, 2000, 35(1-3): 75-86

[17] Harrald, John R.; Stephens, Hugh W.; vanDorp, Johann Rene; A Framework for Sustainable Port Security, *Journal of Homeland Security and Emergency Management,*2004,1(2)

[18] Massimo Paolucci, Roberto Sacile, Antonio Boccalatte, Allocating crude oil supply to port and refinery tanks: a simulation-based decision support system, *Decision Support Systems*, 2002, 33(1): 39-54
[19] E. Palazzi, F. Currò, B. Fabiano, Simplified Modelling for Risk Assessment of Hydrocarbon Spills in Port Area, *Process Safety and Environmental Protection*, Volume 82, Issue 6, November 2004, 412-420
[20] Malcolm L. Spaulding, J. Craig Swanson, Kathy Jayko, Nicole Whittier, An LNG release, transport, and fate model system for marine spills, *Journal of Hazardous Materials*, 2007, 140(3): 488-503
[21] K. F. Cheung, A. C. Phadke, Y. Wei, R. Rojas, Y. J. -M. Douyere, C. D. Martino, S. H. Houston, P. L. -F. Liu, P. J. Lynett, N. Dodd, S. Liao, E. Nakazaki, Modeling of storm-induced coastal flooding for emergency management, *Ocean Engineering*, 2003, 30(11): 1353-1386
[22] Byung Ho Choi, Efim Pelinovsky, Dong Chul Kim, Kyeong Ok Kim, Kyung Hwan Kim, Three-dimensional simulation of the 1983 central East (Japan) Sea earthquake tsunami at the Imwon Port (Korea), *Ocean Engineering*, 2008, 35(14-15):1545-1559
[23] Jim W. Hall, Ian C. Meadowcroft, E. Mark Lee, Pieter H. A. J. M. van Gelder, Stochastic simulation of episodic soft coastal cliff recession, *Coastal Engineering*, 2002, 46(3): 159-174
[24] Huachen Pan, Esa Eranti, Flow and heat transfer simulations for the design of the Helsinki Vuosaari harbour ice control system, *Cold Regions Science and Technology*, 2009, 55(3): 304-310
[25] Rui C. Guimarães, Brian G. Kingsman, Simulation-optimisation: The method and its application in the analysis of grain terminal operations, *European Journal of Operational Research*, 1989, 41(1): 44-53
[26] Josefa Z. Hernández, Juan M. Serrano, Knowledge-based models for emergency management systems, *Expert Systems with Applications*, 2001, 20(2):173-186
[27] Wan Hu, Yang Qing, Yu Ming-hui, Fei Qi, Grid-based platform for disaster response plan simulation over Internet, *Simulation Modelling Practice and Theory*, 2008, 16(3):379-386
[28] Feng Xu, Xuping Chen, Aizhu Ren, Xinzheng Lu, Earthquake Disaster Simulation for an Urban Area, with GIS, CAD, FEA, and VR Integration, *Tsinghua Science & Technology*, 2008, 13:311-316
[29] M. Kulawiak, A. Prospathopoulos, L. Perivoliotis, M. łuba, S. Kioroglou, A. Stepnowski, Interactive visualization of marine pollution monitoring and forecasting data via a Web-based GIS,*Computers & Geosciences*, 2010, In Press, Corrected Proof, Available online
[30] Ameya Shendarkar, Karthik Vasudevan, Seungho Lee, Young-Jun Son, Crowd simulation for emergency response using BDI agents based on immersive virtual reality, *Simulation Modelling Practice and Theory*, 2008, 16(9):1415-1429
[31] Songbai Cheng, Guohua Chen, Qingguang Chen, Xueying Xiao, Research on 3D dynamic visualization simulation system of toxic gas diffusion based on virtual reality technology, *Process Safety and Environmental Protection*, 2009, 87(3): 175-183
[32] Albert Douma, Marco Schutten, Peter Schuur, Waiting profiles: An efficient protocol for enabling distributed planning of container barge rotations along terminals in the port of Rotterdam, *Transportation Research Part C*: Emerging Technologies, 2009, 17(2):133-148
[33] Alan C. Cowdale, Simulation Modeling Support of Emergency Fire-Fighting in Norfolk, *Proceedings of the 2003 Winter Simulation Conference*, 2003, 1707-171

Session E2: Business Process Management for Logistics

·**Day2: Sep. 16 (Thu.)**

·**Time: 09:00 - 10:20**

·**Chair: Jaehyun Kong**

·**Room: Azalea & Lilac, 5F**

TOWARDS PROFITABLE SUPPLY CHAINS BY MEANS OF THE INNOVATION PROFIT MODEL

Marcel W. Ludema[1*] and Jan J.T. de Vlieger[2]

[1]Faculty or Technology, Policy & Management
Delft University of Technology
Jaffalaan 5
Delft, The Netherlands
e-mail: m.w.ludema@tudelft.nl

[2]Innovation Dragons Consultancy
Gerrit Rietveldstraat 3
Rotterdam, The Netherlands
e-mail: jjdevlieger@innovationdragons.nl

Abstract: The well known Strategic Profit Model (SPM) helps company's management to determine and simulate the impact of business decisions on key financial parameters, including those related to the supply chain. In this paper the opportunities to incorporate innovation-related issues in the SPM are explored, resulting in the Innovation Profit Model (IPM). In developing the model, particularities of the innovation process are taken in to account and integrated, including the Service/Product Life Cycle, portfolio management, and the internal and external innovation process. The IPM can be used as a toolbox for companies to understand and manage their innovation potential. Practical use of the IPM is demonstrated in a case study of a company that designs, produces and sells protected clothing. Based on the results, a set of general recommendations for the use and the future development of the IPM are given to make it even more feasible in the use of advising businesses to gain sustainable profit by innovation.

1. INTRODUCTION

Innovation management is seen as an important issue in the sustainability of businesses. A quick review done with a limited set of companies and a review of innovation experiences of companies indicates that many showstoppers to realize fruitful innovations exist. These include: (1) lack of resources and/or organizational shortcomings, for example resulting in a discontinuous, ad-hoc innovation process; (2) not having the right attitude or not having enough targeted experience towards innovation; for example it is often indicated that innovation is not needed in view of strong current business results; (3) not having a clear and brought vision about the functioning of the company in its environment; this includes for example companies claiming that they have explored all the potential routes already and see no opportunities for more or alternative innovation options.

In Medium and Small Business innovation is often evoked by demand of a customer. Still companies do not see innovation as the driver for daily business that relates to all aspects of the business, and how it helps to build a sustainable portfolio for targeting different markets with different services and products. Better understanding of innovation management and how it should be linked to strategy, supply chain management (procurement, production processes and sales) next to marketing, and research and development is advised for companies to make full use of the potentials of innovation to create healthy and sustainable businesses.

Companies at the same time do not see the relationships of their service/product markets with the existing company fulfillment structure let alone the fulfillment structure of the companies that are a member of the same value network. If innovation is seen as something creative, unattainable and 'nice to have not need have' or is left to the creative potentials of a general director, owner, creative marketing manager or the whiz-kid in R&D, the business misses out on opportunities to excel and prosper from its potentials.

Several methods and tools applied in others dimensions of business can be helpful in identifying the role of innovation and how it relates to the supply chain function and the overall business. The Strategic Profit Model, the concept of Service/Product Life-Cycle the service/product portfolio concept, the value network concept and the Balanced Scorecard can be interrelated in such a manner that a toolbox for understanding the innovation potentials of companies can be established. This paper shows components and the current composition of the toolbox; Innovation Profit Model.

The outline of the paper is as follows. In the next section some insight in the service/product design in (industrial) innovation processes is given. This includes topic like the ingredients and structure of the innovation process and the role of the service/product life cycle and the service/product portfolios. Next the Strategic Profit Model is explained and how it could be utilized as the basis for a toolbox for management to steer innovation management. Next a first conceptual design of the Innovation Profit Model is elaborated upon. Before conclusions are drawn and recommendations for further research are given, a case study demonstrates the use of the Innovation Profit Model.

2. SOME NOTIONS ON INNOVATION MANAGEMENT

2.1 The Innovation Process

Currently we still live in the age of innovation. After the focus on efficiencies in the 1950s and 1960s, the focus on quality in the 1970s and 1980s, the focus on flexibility in 1980s and 1990s the focus in the 2000s and probably the 2010s is on innovation, (Bolwijn and Klumpe, 1990). Industry, even paced as a result of the recession, is increasingly applying new technology to new services, products and processes. In most organizations technologies from a wide range of disciplines has become an important foundation for change and innovation. Innovation makes it possible to either become more competitive in the same market or to open up new markets. It sets opportunities for doing things differently, to bring forward services and products at lower cost by means of driving out wastes in fulfillment processes, to shorten the time-to-market of new services and products, to add functionalities to services and products better in line with the needs of customers and to establish the right quality for the right purpose.

Innovation is connecting the business side with the technology side within companies, branches, industries and countries. Non-business focused changes in technology can be seen as inventions that may but are not necessary tuned to new, or renewed commercial activities. "Innovation" is a term defined by Joseph Schumpeter (1939) as "the commercialization of new combination". It can be seen as a set of activities companies can conduct: (1) the introduction of new materials and/or new components and/or new technologies; (2) the introduction of the application of new processes and/or ways of working; (3) the opening and/or targeting of new markets; (4) the introduction of new organizational forms.

For companies innovation is necessary to survive in the market and helps to build up sustainable advantage. Doing things differently can be accomplished by means of buying companies that already are doing things differently and incorporate these things into the business processes or to create "doing things differently" within the company itself by means of a structured process.

Innovation is not at all easy to manage and a targeted organizational structure and next to that structured procedures should be in place to guide the innovation process. Permanent innovation is necessary to keep or gain a sustainable competitive advantage. To be able to do this acquiring, maintaining and managing relevant knowledge is needed. Knowledge management is crucial for innovation management. Those companies that have insight in innovation opportunities together with possibilities to utilize superior skills in applying this insight will be more successful than the competition. By means of innovation, companies increase net value for customers and enforce the relationship with these customers. The innovation process is a trajectory aimed at revisiting Technology, Application, Markets and Organizations till a satisfactory outcome is reached, (Janszen, 2000).

Businesses are part of supply chains that are part of value networks that connect through industrial clusters and markets. Companies are part of more than one value network and sometimes form temporally or permanent alliances as part of buying or developing market power/opportunities, services/products or specific knowledge for their new or continued business. New Product Development (NPD) and New Business Development (NBD) for services and products that contain more functions and thus are more complicated because of the application of more and other technologies than before takes place in complex environments in which many stakeholders have a role, more actors have to be taken into account, and where many issues have to be addressed. Cross functional teams with changing compositions during the innovation process are often in place.

In more branches and industries customers are making higher demands upon suppliers and their products. In response to these constraints several new management approaches are evolving. The first hurdle to take for companies that did not evolve towards the age of innovation is to make sure that they re-engineer their organization towards an innovative organization. After building the innovative organization that promotes innovation it is time to reach out to suppliers and customers. One of the most important changes in management is the notion that competition is less company focused and more supply chain focused. Demand/supply chains compete with each other rather than companies. Supply chain integration, partnering and collaboration strategies are becoming more important. This may be in the form of: co-makership, co-design, strategic alliances, networks, hybrid organizations, virtual organizations, collaborative and concurrent engineering, and parallel product development that facilitate co-development. The innovation and knowledge held throughout the entire supply chain is extremely valuable to the innovation and product development process (Bruce and Cooper, 2000). Innovation is thus not a 'stand alone' function in the business but is closely linked to, or integrated into, business strategy and development, internal and external organization, and ultimately profit/loss and the balance sheet.

3.2 Structuring the innovation process

The innovation process is not a linear process but often explained as one. A structured description of the innovation process helps to identify the right milestones and deliverables and at the same time helps to understand how organizational parts of the company or demand/supply chain might be involved to successfully execute the innovation process. A formalized innovation process that can be followed contains the following phases, (Roozenburg and Eekels, 1995):

1. Stating the idea for a new business activity; the service/product-market plan should follow or be aligned with the strategy of the company, the current and anticipated position of the company on the market (differentiator, cost leader or being in a niche market), (Porter, 1996). A feasible balance between technology push and market pull should be understood and elaborate upon, and it should be clear what the assignment for service/product development that underlines the innovation process is in terms of product idea, expected volumes, price range and product positioning.
2. It is necessary to reframe the companies service/product development policies based on goals and strategies. Every company should have as policy regarding their goals and strategy towards the markets and how services and products relate to this policy. The policy gives direction and continuity to the behavior of the company in the market it operates in and direction to the innovation process leading to new or renewed services/products to sustain or redirect their market position. Policies touch on issues such as: size of the company in the long run, the expected financial status and how it related to the new business activity, chosen strategies in respect to the portfolio of service/product market combinations the company wants to offer now and over time, the availability of dedicated technology; the possible alliances in respect to scale & knowledge.
3. Analyzing and evaluating ideas on their feasibility and selecting ideas for further development is needed. Based on customer needs, market information, and a SWOT analysis of the company and its environment, a first selection of the most feasible ideas for innovation of services, products and processes should be made. After the analysis the ideas should be compared and scored and the most feasible ideas should be selected. Finally it should be decided for which idea(s) a formal New Service/Product Development process is going to be started.
4. Formal (strict) development of products in a New Service/Product Development process is necessary. This process has two important perspectives the technical and the commercial that both have to be developed in close relation to each other. The technical development rounds with clear milestones, deliverables and decision points, so different stakeholders and disciplines will have to participate till a satisfactory result is realized. A commercial development process is necessary to calculate the potential profit, revenue and cost, based on a marketing plan specifying the service/product in relation to 'place, promotion and price'.
5. Realizing the service/product using the infrastructure available to the company and its partners. Factors to take into account are the availability of means to produce / offer services in relation to an acceptable 'time to market' and synergies with the production of the company's current services/products. Understanding what the customer-order decoupling point will be and how it relates to the overall functionality including the control mechanisms of the service/product production process is essential to design an effective and efficient a fulfillment structure.

In customer-driven industries companies face several challenges that are fed by: intensified competition, more demanding customer (for broader choices and rapidly changing tastes and preferences), relentless pressure for efficiency, waste reduction standardization, modularization, and more off the shelf items), efficiency, and volume (scale) versus variety trade-offs, as well as adaptability to several demand dynamics. More and more industries are becoming pull-based systems to (re)plan coordinate production and distribution with actual and real-time customer demand.

3.3 Service/product design

Services (not tangible products) and products (tangible) are 'artifacts' conceived, produced or developed, transacted and used or conceived by people because of their properties and the functions they perform. Service/product design is the process of devising and laying down the plans that are needed for the development of a service or the manufacturing of a product. Designs can be described with the use of several cognitive aids like texts, schemata's, technical drawings and many other aids. Designing can be supported by a wide variety of methods.

Typical the following five stages within the life of a service/product that differ from each other in sales and profitability can be distinguished: (1) development, (2) introduction, (3) growth, (4) maturity, and (5) decline. The basic idea behind the product life cycle is to show that the pattern of the sales history of any service/product reminds of the life cycle of an organism. No (industrial) service/product has eternal life, sooner or later it will disappear from the market, either it is replaced by an improved alternative, or the need for it has disappeared. It is clear that companies are fighting a constant battle for survival with their services/products. Keeping the set of new and available services/products competitive in respect to changing market demands at a reasonable price is what running a business all is about. Because over time competitive services/products will be offered by many others, prices will have to drop and over time available alternatives to produce cheaper (cheap labor, cheap raw materials, depreciated plant) are diminishing.

A well balanced service/product range of a company consists of services/products that are at various stages of their followed and anticipated Service/Product Life Cycle. There will be newly introduced services/products, growing services/products, mature services/products and even services/products in decay in the service/product portfolio. A continuous adjustment of the range is necessary. As obsolete services/products are taken out of service/production new services/products ought to be added. In progress the service/product portfolio of a company will change. From this principle it is clear that companies should force to them selves to develop new business activities that replace current activities. The potential for a company to develop new services/products determines the company's life cycle.

3.4 Requirements for a toolbox to manage innovation

In the previous section several concepts that tune service/product innovation to operate a business were discussed. Several shortcomings of properly addressing innovation temporally, the role and monitoring the service/product life cycle of the and consequently understanding the relation of the product portfolio in respect to the service/product portfolio where elaborated upon. Furthermore a workable business dashboard 'the Strategic Profit Model" was explained. It was suggested that an extension and/or addition of this model would provide a feasible model for keeping track of profitability in relation to the company's service/product portfolio and via the service/product life cycle be able to show how it relates back to the role of innovation management and the potentials of frequently develop new services and products tuned to a aligned and cost aware fulfillment structure. Based on the addressed issues the requirements of a toolbox to manage innovation are:

1. The ability to show the relationship between portfolio renewals in services, products and processes;
2. The ability to show how estimated profits (and costs and revenues) relate to these portfolio renewals based on scenario analysis and simulation;
3. The ability to understand the relation between potential saving on current services and products in the portfolio to streamline the fulfillment processes as part of the company's business processes and how these relate to necessary investments in new service/product and necessary fulfillment structures.
4. The ability to relate the business potentials for innovation of a company to other companies in the value network. Basically to understand how the use of the Innovation Profit Model for a company relates to the Innovation Profit Models of other companies in the same value network.
5. Provide the potential to understand all above from a dynamic standpoint and to be able to show and analyze how the solution area relates to changes in variables available.

4. THE STRATEGIC PROFIT MODEL & INNOVATION MANAGEMENT

4.1 The Strategic Profit Model

The Strategic Profit model (SPM) is the formalized version of the DuPont Model, a model developed at DuPont to help managers understand how changes in operations impact shareholder value (Shapiro and Kirpalani, 1984). The SPM shows the tradeoff between the net profit margin and asset turnover as it can be used to analyze the competitive position of a company against it peers or the competitive position of a service/products in a company's service/product portfolio. The SPM demonstrates that Return on Net Worth, a key indicator of shareholder value, is determined by Net Profit, Asset Turnover and Financial Leverage, all of which can be controlled by company management. The model is based on two equations: (1) Return on Assets (ROA) = Profit Margin times Asset Turnover, and (2) Return on Net Worth (RONW) = ROA times Equity Multiplier. ROA is a measure for how well a company manages to make a profit on its assets; RONW is a measure for how well a company uses shareholder investment. The model takes the form as indicated in figure 1.

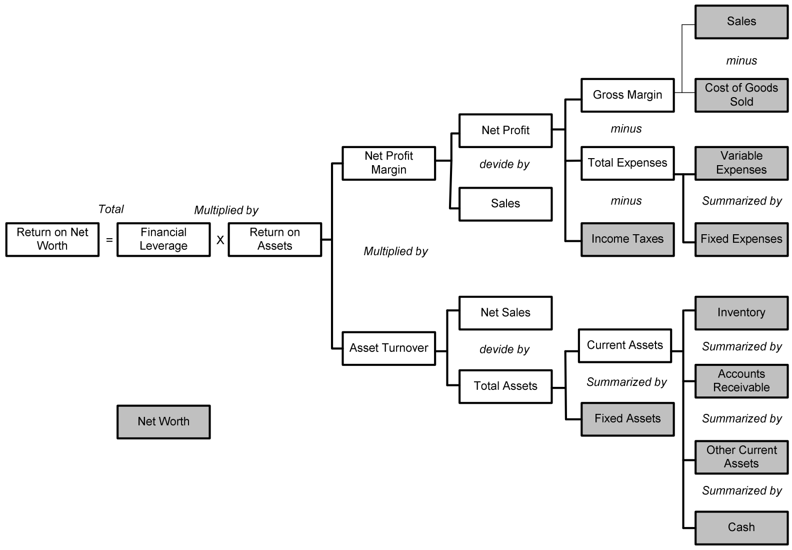


Figure 1. The Strategic Profit Model (based on Lambert and Stock, 1993)

The figure indicates how RONW and ROA are determined based on data from the company's balance sheet and profit & loss statement. It follows from the model that to increase RONW and ROA, companies need to work on increasing/decreasing (a combination of) the components of the model, i.e. increase sales, decrease costs/taxes/expenses, decrease inventory and accounts receivable, decrease assets, and increase financial leverage.

By simulating the effect of particular measures, company management can determine which are the most effective and realistic measures to obtain a particular improvement in RONW and ROA. Stapleton et al (2002) have followed this approach to measure and compare the performance of several sportswear companies, with a particular focus on the logistics function. Their work effectively demonstrates how the Strategic Profit Model can provide company management with actionable recommendations relating to company operations in order to improve shareholder value.

4.2 Applying the Strategic Profit Model to innovation management

For the following reasons, the Strategic Profit Model (SPM) is considered as a useful and relevant model to guide and support innovation management decisions. The SPM directly ties into the financial performance of a company. As discussed before, innovation plays a vital role in securing a company's profit and sustainability. As motivated in paragraph 3.4, our take on innovation is it requires companies to tailor the product and fulfillment structure to the service/product life cycle stage. All these decisions have implications for revenue/cost, cash flow and asset turnover, all of which are reflected in, and can be simulated using, the SPM. Notwithstanding its relevance, applying the SPM to innovation decisions is not the full answer to successfully managing innovation from a financial and product life cycle perspective. Here three shortcomings are listed that, when addressed, can help the SPM to evolve into a more suitable and more comprehensive tool to manage innovation.

1) Typically, companies provide a range of services/products each of which are in a particular service/product life cycle stage. It is vital that the model is able to cope with this service/product portfolio in the sense that next to the company as a whole, its individual services/products can be evaluated. For each service/product market combination a SPM could be made and analyzed in relation to the particular product's life cycle stage.
2) It is vital for companies to realize that in order to remain profitable, innovation relates to investing in new services/products as well as to disinvesting in existing services/products and increasing efficiency, enabling the company to fund new service/product development. This concept is here referred to as 'two-sided innovation'.
3) As argued by Porter (1996), company strategy is to be closely aligned with a company's services/products, the fulfillment structure used, and the markets it focuses on, now and in future. Touching on all of these aspects, innovation should thus be closely aligned with the company's strategic direction which services/products, markets and fulfillment structure (including organization of the supply chain and coordination with value network partner) best fit its strategy, now and in future.

5. INTRODUCING THE INNOVATION PROFIT MODEL

5.1 Principles and elements of the model

In this section the well known Strategic Profit Model is adapted towards the Innovation Profit Model (IPM) to give insight in the possible effects of the streamlining of business processes and creating new service and products. This model gives insight in the contribution of increasing the profit of individual companies and the aligned value chain they are part of. The IPM shows the relationship between portfolio renewals in services, products and processes and how they are linked to costs and revenues. In essence how profit can be influenced by governing the service/product portfolio and the fulfillment processes that are needed to bring them forward. It is particularly interesting what the possibilities are to innovate without additional budgets for new service/product development. The main principles of the Innovation Profit Model are: (1) An integrated analysis of most of the company issues and their close connection to innovation (innovation as the main reason for running a business, achieving innovation by means of new or renewed services/products/processes, considering that innovation can be renewal and cost saving at the same time, viewing innovation from a dynamic perspective, simulating the effect of innovation closely with company strategy). (2) Taking a Value Network perspective rather than an isolated company perspective: taking an inside-Out viewpoint of innovation looking at the strengths of the whole value network a company is part, understanding the needs and capabilities not only of the company's own customers and suppliers but also taking those of the customers' customers and suppliers' suppliers and into consideration, working towards alignment of processes and building the right partnerships to expel potential innovation opportunities taking a shared revenue and cost perspective. (3) Govern and manage a balanced product-service portfolio in respect to the underlying fulfillment infrastructure and its processes by mapping the current position of services and products on their respective product-life cycle. (4) Building scenario's for the future positions of respective product-life cycle, developing strategies for renewal and de-investing in business activities based, making choices of developed strategies based on detailed insight in the needs of customers and other stakeholders in the value network. Normally a detailed analysis of the portfolio of service/product life cycle shows:

- A balanced portfolio consists of a good spread of services/products over the service/product life cycle;
- Because markets are dynamic and the relative position of services and products and the profits, sales-revenue and costs that related to them change continues analysis and mapping is necessary;
- Companies are the result of innovation and its success can be measured by means of the results of the cumulative portfolio of service/products life cycles;
- Companies should develop the potential to adapt to changing market situations and opportunities and in that respect the portfolio establishes itself as the relative company life cycle. The relative position of companies and their competitors can be plotted on the product-life cycle of certain product families.

5.2 Structure of the model

Combining the above, the Innovation Profit Model proposed is a tool to support management in taking innovation-related decisions based on multiple criteria. The requirements determine the strategic fit of an existing/new service/product, the indicators determine the stage in the service/product life cycle, the (expected) contribution to Profit Margin and the (expected) asset turnover rate, the diagnostics determine the effectiveness and efficiency of the related fulfillment structure.

Based on an assessment of the above criteria for individual service/products in relation to the other services/product in the portfolio, company management can decide on the most suitable action to be taken. These actions can include:

1. Investing, or speeding up investments, in the service/product;
2. Disinvest, outsource or sell the service/product line;
3. Increasing efficiency of the fulfillment structure;
4. Increasing or reducing service/product variety and complexity

5.3 Case Study: Safe and Rough Clothing BV

The Innovation Profit Model was validated in an extended case study of a company, SRC bv, which designs, produces and sell protected clothing. The company has three business units that focus on the following clients groups: fire departments, oil/gas/chemical companies and maritime industry. The Oil/gas/chemical business unit is discussed that account for the majority of revenue, 35 million euro annually. For SRC, the Oil/gas/chemicals business unit shows a low profit margin and asset turnover rate compared to the other business units in the company and a selection of competitors.

Table 1. Product comparison analysis

	Sales	Profit margin	Asset turnover	OA
Product A	11,943,800	2.9%	1.64	.8%
Product B	5,983,125	4.4%	2.26	.9%
Product C	3,283,077	2.2%	1.54	.4%
Product D	7,345,094	0.9%	1.49	.3%
Product E	4,331,961	9.4%	2.40	2.6%
Product F	1,987,463	-6.2%	2.44	15.1%
Total	**34,874,520**			
Mean	**5,812,420**	**3.0%**	**1. 71**	**.2%**

Looking into the business unit's product portfolio, 6 products/product lines can be identified that account for the majority of total revenue. Currently no products are in the NPD pipeline. Following a detailed inventory of revenues, costs and the contribution to these of each of the 6 products can be identified, see table 1. Looking at the product life cycle, the picture as in figure 2 emerges (based on qualitative assessment following discussions with company management and external market experts).

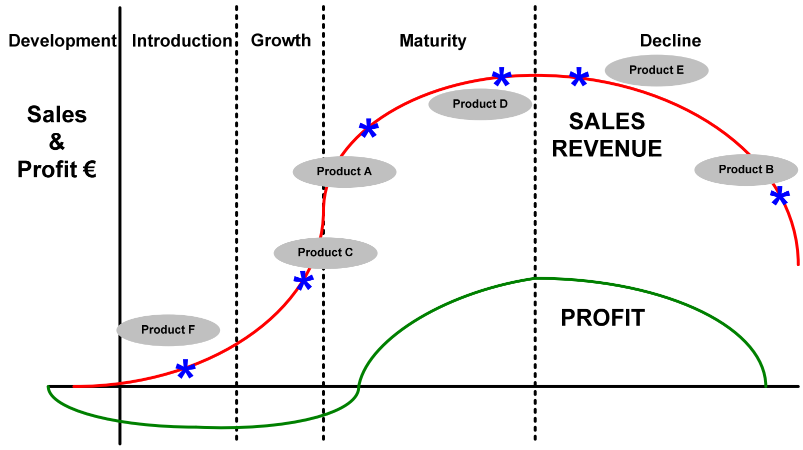


Figure 2. SRC stage of products portfolio in the product life cycle

First focusing on product F, in terms of contribution to sales it is a relatively small product yet it is only in the introduction stage. It fits well with the company structure and shows good future growth prospects according to a market study conducted the preceding year. It is also encouraging to see a high asset turnover ratio, which can be attributed to low inventory levels of this product, which is very much in demand. The company is recommended to continue investing in this product which can be an important product once mature. To limit losses in the short term, SRC is recommended to explore opportunities to raise prices, which the market is expected to absorb without a significant drop in demand.

Product A is the company's largest product in terms of sales, and shows an average profit margin in combination with a lagging asset turnover ratio. The product is still is the maturity stage. In view of technological developments in light weight materials, this product is expected to enter the decline stage relatively soon. A simulation of opportunities to at least maintain the profit margin and increase asset turnover ratio suggests that the optimal solution for this product is to reduce the range of variants offered, resulting in lower costs of goods produced, lower inventory levels, and at the same time simplify the order intake procedure which also has a positive impact on accounts receivable.

Simulation shows that the combination of the above and other measures would raise the company's overall profit margin to 6.0% and its asset turnover ratio to 2.2 which brings it in line with its key competitors. In view of considerable savings made, no additional budget is required to fund the further development of products in the early stage of their life cycle, while a modest but sufficient budget will become available for starting to fill the new product development pipeline, thus catering for the company's continuity in the longer run.

6.CONCLUSIONS AND RECOMMENDATIONS

Innovation seems to be the most important driver for successful and sustainable business. Currently many showstoppers to innovation exits leaving full potentials of companies untouched. Innovation should be made an approach linking company strategy via research and development, marketing and sales and logistics next to finance. It calls for a structured innovation development and life cycle management approach. Clear models for these though often decoupled or loosely coupled are available. Shown is that the service/product life cycle concept together with service/product portfolio governance and using a clear business dashboard as the Strategic Profit Model to be able to related service/product life to attainable profit levels and at the same time viewing this from a value network perspective gives the right potentials for running an innovative and sustainable business. The innovation Profit Model is the toolbox that has the potentials to offer this feasibility and at the same time offers the possibility to show the dynamic character of possible innovation related business changes that have to make or can be made by the management of companies.

The application performance of the Innovation Profit Model can be improved in several areas. This requires applying it to more realistic situations. Better comprehension of the dependency of variables in the Innovation Profit Model. Decisions to lower inventories or stepping up NPD investments will impact on other factors such as sales, debts, etc. The key relationships that potentially can distort the model if not taken into account should be incorporated. A more structured set of decision criteria, e.g. in the form of decision trees that point towards particular 'default' decisions or actions needs to be taken.

7.REFERENCES

Bolwijn, Piet and Klumpe, Ted (1990). Manufacturing in the 1990s: productivity flexibility and innovation. *Long Range Planning, 23 (4):* 44-57.

Janszen, Felix (2000). *The age of Innovation; making business creativity a competence, not a coincidence,* Financial Times Prentice Hall, London.

Lambert, Douglas M. and Stock, James R (1993). *Strategic Logistics Management, 3rd Edition*, Boston, Mass: Irwin-McGraw Hill, pp. 42-215.

Margaret, Bruce and Cooper, Rachel (2000). *Creative Product Design; a practical guide to requirements capture management.* John Wiley & Sons, Chichester, UK.

Porter, M.E. (1996), What is Strategy, *Harvard Business Review*, *Nov/Dec 1996.*

Roozenburg, N.F.M. and Eekels, J. (1995). *Product Design: Fundamentals and Methods.* John Wiley & Sons, Chichester, UK.

Schumpeter, Joseph A. (1939). *Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process.* McGraw-Hill, New York, USA.

Shapiro, Stanley J. and Kirpalani, V.H. (1984). *Marketing Effectiveness: Insights from Accounting and Finance.* Allyn and Bacon Inc., Boston, Mass, USA.

Stapleton, Drew, Hanna, Joe B., Yagla, Steve, Johnson, Jay and Markussen, Dan (2002). Measuring Logistics Performance Using the Strategic Profit Model. *International Journal of Logistics Management 2002 Issue 13-1.*

A CONTEXT-BASED FRAMEWORK FOR FLEXIBLE CONTROL OF BUSINESS PROCESS IN LOGISTICS ENVIRONMENTS

Jaehyung Kong[1] and Hyerim Bae [1*]

[1]Department of Industrial Engineering
Pusan National University
Jangjun-dong, Geumjung-gu
Busan, Korea
e-mail: kongjaehyun@gmail.com
hr bae@pusan.ac.kr

1.INTRODUCTION

In logistics environments, it is important to take into account situation when a logistics process flows over heterogeneous partner. Especially, due to dynamism of logistics process, information system should have ability to control logistics process in a flexible way according to current situation. Through emerging ubiquitous technology including RF ID, gathering information on context-awareness became possible. In this paper, we suggest a context-aware framework based on RFID technology for business process management. Using RFID, it is possible to collect information about location of workers and object, time of arrival and departure, sequence of event and so on. We develop inference rules for context awareness through organizing business constraint, and construct system architecture for context-aware business process management. And we suggest a method to deal with RFID data for context-awareness. We introduce inter-workflow patterns and method for context-awareness with inter-workflow patterns using RFID technology.

2. INTER-WORKFLOW PATTERNS

In logistic environments, a process, in that it manages the flow of materials among partners, inherently involves more than on organization. Traditional business process management cannot be used for the entire flow, since it lacks the ability to manage interactions among partners. Then, Bae (2010) propose inter-workflow patterns that represents the relations among separate processes. He introduces inter-workflow patterns that have been discovered through our several years of research on logistic process modeling. There are seven inter-workflow patterns.

- Chained Service Model (CSM) – Once a process completes, another process in triggered and commence its execution.
- Nested Sub-Process Model (NSpM) – A process becomes a unit task of another process, forming a parent-child relation.
- Parallel Synchronization Model (PSM) – The PSM comprises a relation pattern in which two different processes are executed in parallel and synchronized at a certain point.
- Multiple Instance Model (MIM) – The MIM enables a process to trigger multiple instances of the other process.
- Instance Merge Model (IMM) – Multiple process instances can trigger a single instance, which situation corresponds to the IMM.
- Exception Handling Model (EHM) – A process plays the role of exception handling for another process or task in which the exception takes place.
- Cancellation Propagation Model (CPM) – CPM is the relation pattern between two processes, by which process failure is propagated to the other process.

In this paper, we suggest the methods that help recognize logistic context and coordinate the process by mean of RFID for six workflow pattern exception CPM.

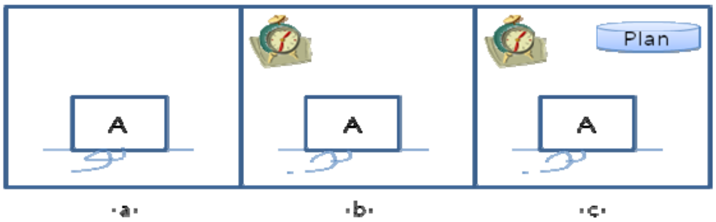


Figure 1. The Level of Context in Logistics Environments

3. CONTEXT AWARENESS

Future pervasive and autonomic service must evolve from a simple model of "context-awareness", in which they access isolated pieces of heterogeneous contextual data, directly in charge of digesting and understanding them (Castelli *et al.*, 2008). However, it is very difficult to define contextual data in logistic environments. Especially, Using RFID(radio frequency identification) technology which is propagated in logistic environments, information which help to recognize context is limited. In this paper, we suggest the method for context-aware in logistics environment by means of RFID technology.

Special situation in logistics environments can be recognized based on information of location, time and planning. Through this information of workers and products, we can recognize situation with completion, waiting, arrival, departure, delay and so on. In this paper, we categorize context in logistics into three levels, as shown in Figure 1.

First, we recognize the location of workers and products using RFID. That is to know existence of specific workers or products then we can perceive arrival or departure of products and start and end of tasks. When some processes are completed in one part of logistics, products leave to succeeding part and eliminated. RFID readers catch disappearance of product, then movement of product is realized. So workers in next process can recognize proceeding job by means of information systems, they will prepare next task. Secondly, we use real-time information in order to check the staying time of products. RFID reader frequently updates the data and we can see how long it remains. Through waiting time information we can coordinate the process on the inter-workflow pattern. Finally, we can use planning data in the information system or specific products data. System reads the product ID through RFID tags and watch for information of planning and specific products information. When the exception occurs, process is progressed based on this information.

4. COORDINATION OF LOGISTICS PROCESS

In logistics business process is progressed through movement of products, in many cases. In the case of using RFID technology, when a certain reader detects the specific product tag data or detected product is disappeared, we can know the departure of product or arrival respectively. Then we recognize the end of processes or start. In this paper, we suggest the methods for connection of processes based on inter-workflow pattern and RFID technology. We don't consider scheduling or worker assignment problem.

4.1 Simple Process Coordination

We consider very simple cases when the preceding process is completed and succeeding process is started. The traveling time is existed between processes. When product is arrived to the succeeding process workspace, if process is ready-condition then process is started immediately. However, if process is not ready-condition(e.g. another process instance occupy workers) then process must be waited. If departure sequence of process instance is deferent from arrival sequence then process ready-condition must be changed.

Table 1. Simple Process Coordination

Case	Context	Action	Example of Rules
1 No	rmal	Progress	-
2	Not Available	Call Rule 1	Wait
3	Out of sequence	Call Rule 2	Prepare for another task

4.2 Waiting Process Coordination

On s ome occasion, waiting t ime is n eeded. If sy nchronization is essential , s ome p rocesses wait for another process completion. Waiting time is minimized, static or planned on the purpose. Expected context and action is as following.

Table 2. Waiting Process Coordination

Case	**Context**	**Action**	**Example of Rules**
4 ur	gent	Start Waiting process immediately	-
5	Static time remaining	Call Rule 3	Wait for limited time
6	Planned time remaining	Call Rule 4	Wait for variable time by planning

4.3 Checking process coordination

Some process change the next activity through inspection. System information is used in these cases.

Table 3. Checking Process Coordination

Case	**Context**	**Action**	**Example of Rules**
7	Requesting for check	Check and Call Rule 5	Normal → Progress Abnormal → Exception Handling

5. CASE STUDY

We show the example case where inter-workflow pattern and RFID te chnology is useful. Figure 2. illustrates the case of logistic process of warehousing and taking of out warehouse.

When produ cts arr ive t o war ehouse by tr ansporter, RFI D rea der i n t he g ate d etects arr ival i nformation. System delivers c ommand f or pre paring lo ading activ ity at the loading doc k. Loading activity is started based o n re al time information of empty space and idle worker(CSM; Case 1, 2, 3). And RFID readers check the information of each product and quantity. If collected information is different from order information then sending bask process is started(EHM; Case 7). If l oading tim e is li mited(e.g. refrigerated gro cery products) t hen we adjust worker's sch edule in order t o m atch static waiting time(NSpM; Case 5). When s toring activity is started, updated i nformation is deliv ered imm ediately in t he inventory management process(PSM; Case 4). In the taking out of warehouse process, if various products are t ransferred using a truck, transportation process is started when every taking out processes are completed(MIM & IMM; Case 6).

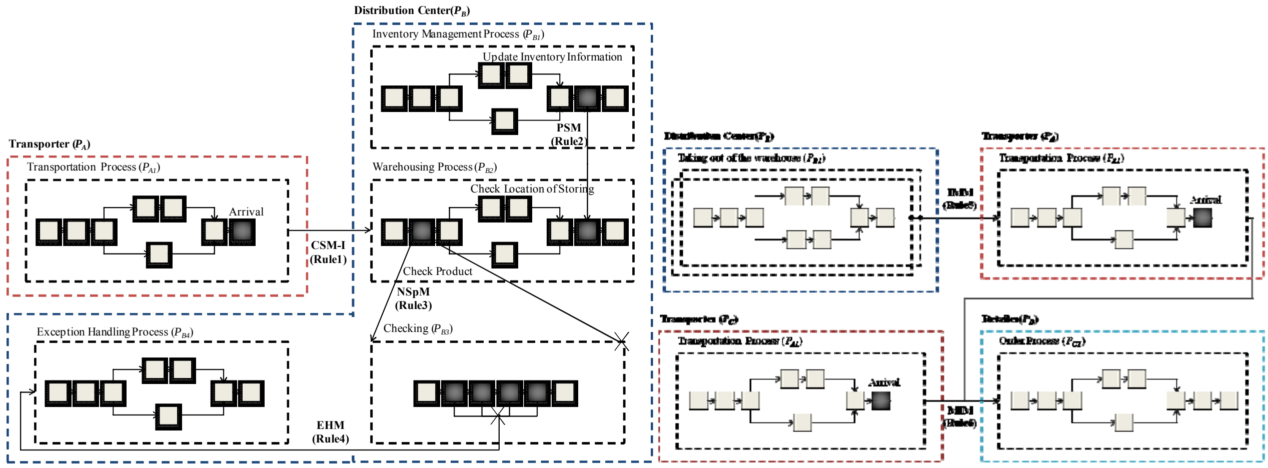


Figure 2. The Example Process of Logistics

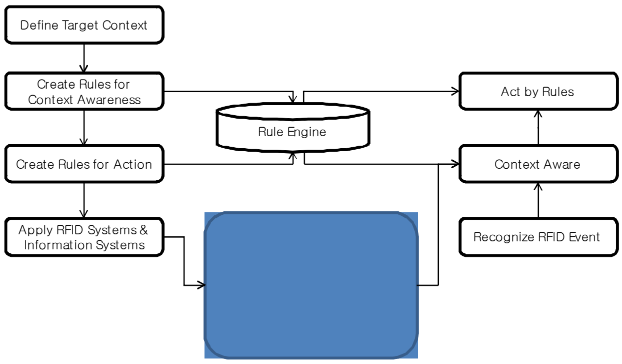


Figure 3. System Architectures

6. SYSTEM ARCHITECTURES

System Architecture is illustrated at Figure 3. First of all, we define target context for each processes in logistic environments. And we create rules for context awareness and action in order to use real time data base on RFID technology. These rules are stored and managed in the rule engine. Small agent is distributed to business information system and RFID system to invoke service in real time.

7. RELATED WORKS

Previous researches have been continued for context-awareness. However, many studies are limited to communicating the systems or integrating information collected by sensors (Castelli *et al.*, 2008). In logistic environment, RFID technology is very useful tools for context awareness. However, the researches for context-awareness using RFID are few except studies about complex event processing(Zang *et al.*, 2008; Wang *et al.*, 2009).

8.CONCLUSIONS

We suggest a method to deal with RFID data for context-awareness in logistic environments. In this paper, we introduce the inter-workflow pattern and method for context-awareness with inter-workflow pattern using RFID technology. We expect that our approach enables business processes to be controlled in a more flexible way and thus we can improve process efficiency and avoid erroneous execution by means of context-based technology.

ACKNOWLEDGEMENTS

This work was supported by the Second Stage of Brain Korea 21 Project and by the Grant of the Korean Ministry of Education, Science and Technology(The Regional Core Research Program/Institute of Logistics Information Technology).

9. REFERENCES

Bae., H. (2010). Inter-Workflow Patterns in Logistic Processes. *Handbook of Research on Complex Dynamic Process Management: Techniques for Adaptability in Turbulent Environments.* IGI Global, Business Science Reference, USA.

Castelli, G., Mamei, M. and Zambonelli, F. (2008). Engineering Contextual Knowledge for Autonomic Pervasive Services. *Information and Software Technology*, 50:36-50.

Zang., C., Fan., Y. and Liu., R. (2008). Architecture, Implementation and Application of Complex Event Processing in Enterprise Information Systems Based on RFID. *Information Systems Frontier*, 10:543-553.

Wang., F., Liu., S. and Lin., P. (2009). Complex RFID event processing. *The VLDB Journal*, 18:913-931.

MAINTEANANCE SYSTEM OF MAIL SORTING MACHINE

Jeong-Hyun Park, Ho-Yon Kim, Byung-Chul Cha, Jong-Heung Park

Postal Logistics Technology Department
Electronics and Telecommunications Research Institute (ETRI)
161 Kajong-Dong, Yusong-Ku,
Daejeon, 305-700, Korea
e-mail: jh-park@etri.re.kr

Abstract: This paper describes the problem of existing maintenance system for mail sorting machines, and presents To-Be considerations of maintenance system based on AHP (Analytic Hierarchy Process) analysis with efficiency and economic aspects. There are standard procedure for diagnosis and repair of mail sorting machine as preventive and operational maintenance, roles and missions of maintenance technical support center for upgrade of maintenance skill in this paper. The proposed maintenance system will be adapted for more safety and efficiency maintenance of mail sorting machine.

1. INTRODUCTION

Maintenance through the last several decades, until recently, was a relatively monolithic central function. It was usually staffed for peak activities, and often had excess capacity waiting for the breakdown to occur. With the advent of international competition in the 80 's, many maintenance staffs were cut dramatically, and over several layoffs became smaller than half their original size. These cuts were often made strictly according to either financial rules (non-union companies laid-off the most senior, expensive workers) or seniority rules (union shops left seniority in place). In neither case were skills and experience the major consideration [EFNMS 1999, S.B Peters on 2005&2002, R.D Hedding 2005]. Simultaneous with reducing costs, companies were forced to increase quality, productivity and safety. These efforts focused on the manufacturing unit, looking to reduce variation in product, reduce production bottlenecks, and assure safe work practices. Quality theory told us to define who our customers are and get close to them. Most plants defined operations as the maintenance customer, and in increasing accountability for operating unit managers, gave them more control of the resources. The initial result was a surge in machine's operability, as operations managers directed resources towards equipment problems that had been chronic problems. The craftsmen dedicated to the units felt needed and like they were making a more direct contribution than before as part of a pool. They learned their unit's equipment intimately, and became more proficient and committed to unit performance.

There are 5 – 6 billion mails/Year, 250 million registered mails/Year, and 60 million parcels/Year without commercial home delivery parcels, in Korea. We also have about 150 different sorting machines in 30 mail distribution centers, Korea. So, it's an important to define the maintenance plan of heterogeneous mail sorting machines in distributed area [Korea Post, J-H Park, and J-H Hong 2008].

This paper describes the to-be considerations such as cost analysis of maintenance system, and diagnosis and repair standard procedures for more safety and economy maintenance of mail sorting machines in Korea.

In section II of this paper, we show the installation, operation, and maintenance status such as failure causes and steps of mail sorting machine in Korea. In section III, we describe the to-be considerations which allow the maintenance of mail sorting machine to focus on their strengths rather than their weaknesses on mail distribution centers in Korea. This paper also shows the maintenance system based on AHP (Analytic Hierarchy Process) for safety and efficient maintenance of mail sorting machine, Korea, in section III, and conclude with summary and further study in section IV.

2. MAINTENANCE STATUS OF MAIL SORTING MACHINE

The usability of mail sorting machine depends on the sorting time and the volume of mail. We show the installation number of sorting machine in Korea, and the sorting time and volume of sorting machine in this section.

2.1 Installation and Operation

2.1.1 Installation

There are 21 mail distribution centers, 2 logistics centers, 1 mail switching center, about 3,600 post offices with 517 delivery centers. We have 53 OVIS (OCR (Optical Character Recognition) VCS (Video Coding System) Interface System)/LSM (Letter Sorting Machine) machines for ordinary letter sorting, 22 Plat Sorting Machines for plat, 28 Packet Sorting Machines for packet, 30 Parcel Sorting Machines which has two types (slide shoe and tilt tray) for parcel, 3 Registered Mail Sorting Machines for registered mail, and 4 Delivery Sequence Sorting Machines in delivery center, in Korea. The figure 1 shows the installation number of sorting machine in Korea, at the end of 2007. There are 43 NEC machines such as NAV-AC18 and NAV-D1KR, etc and 10 Toshiba machines such as TT1000 and TT1021. The 22 flat sorting machines are 19 from Elsag, Italy, and 3 from Siemens, Germany. The 30 parcel sorting machines are 11 from Crisplant, Denmark, and 8 from Vanriet, 5 from Beumer, Germany. The 3 registered mail sorting machines are from Ulzi, and 4 delivery sequence sorting machines are from Ulzi and DaeSeong, Korea [Korea Post, J-H Park, J-H Hong, 2008].

Classify of Mail Sorting Machine		Manufacture	Number of Mail Sorting Machine
OVIS/LSM		NEC/Toshiba	59 (OVIS: 55, LSM: 44)
Flat Sorting Machine [FSM]		ELS/SIE	22
Packet Machine [PKM]		BEU/CRP/MAN	20
Flat Packet Machine [FPM]		BEU	8
Parcel Sorting Machine [PSM]	Slide Shoe	CRP, BEU, MAN, Okura	10
	Tilt Tray		20
Registered Sorting Machine [RSM]		Ulzi	3
Delivery Sequence Sorting Machine [DSSM]		Ulzi, DaeSeong	4
Total			146

Figure 1. Installation Number of Mail Sorting Machine [Korea Post 2008]

2.1.2 Operation

The figure 2 shows the sorting volume of OVIS machines at each mail distribution center in Korea. Seoul, East Seoul, Busan, and Daegu have a large sorting volume as the big size mail distribution center, Suwon, EuJeongbu, DaeJeon, AnYang, Bucheon have a medium sorting volume, WonJu, JeonJu, KoYang have a semi-medium, others have small size mail distribution. The figure 3 shows the sorting time of JeonJu as a semi-medium size mail distribution at the end of 2007. JeonJu has total 8,136 mail sorting hours per year by 2 OVIS, a FSM, and a PSM. OVIS which installed on JeonJu has the sorting rate of 30,000 letters per hour, PSM has the sorting rate of 6,000 parcels per hour.

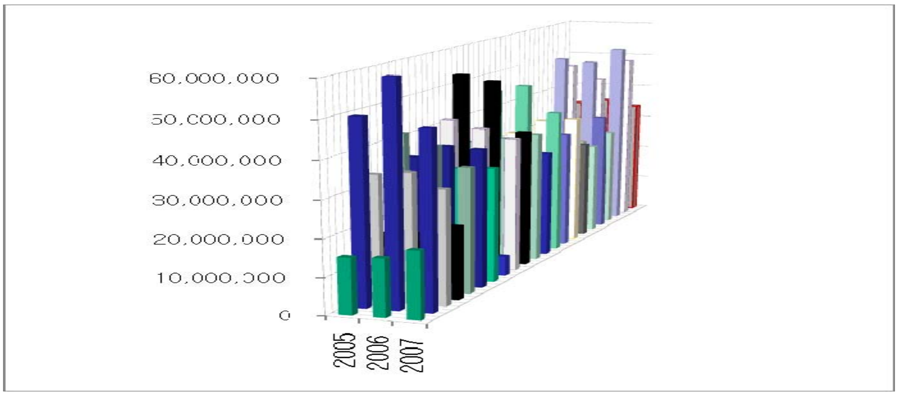


Figure 2. Sorting Volume of OVIS Mail Sorting Machine [Korea Post & J-H Park 2008]

Mail Sorting Machine	Manufacture	Installation Year	Sorting Time		
			Total Sorting Time(Year)	Total Sorting Time(Month)	Total Sorting Time(Day)
OVIS1	Toshiba	2000	1,568	130.6	5
OVIS2	NEC	2003	2,220	185	7.5
FSM	Siemens	2000	1,886	157	6.4
PSM	Mannesmann	2000	2,462	205	8.3
계			8,136	677.6	27.2

Figure 3. Sorting Time of Mail Sorting Machine (JeonJu Mail Distribution Center) [Korea Post & J-H Park 2008]

2.2 Maintenance

In this section, we show failure causes and steps of mail sorting machine [MSM], and find major failure cause and step via statistics analysis of failure data of sorting machine in Korea.

2.2.1 Failu re Cases

We have 7 failure cause of mail sorting machine on current maintenance system such as end of part life cycle [PLC], bad part [operation failure of part, OFP], bad pr oduction of part [manufacturing error of part, MEP], bad use [misuse of part, MUP], bad tre atment [mistreat of part, MT P], cause from other part [COP] , others [O]. The figure 4 shows the statistics data from 2001 to 2007 of failure cause of mail sorting machines. The major failure cause of mail sorting machine based on figure 4 is L CP o f 40 % and OFP of 1 5%. It need s more s ystematic sup port such as stan dard procedu re o f par t, the definition of priority of parts, and part list.

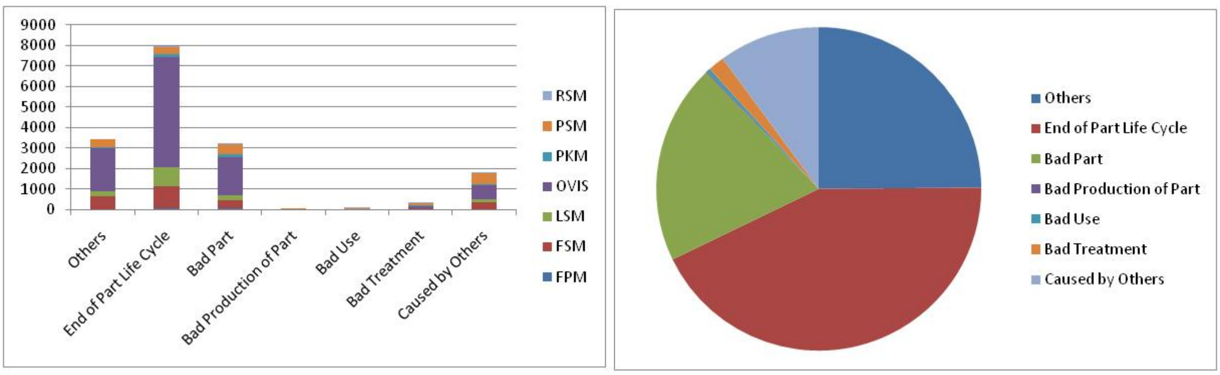
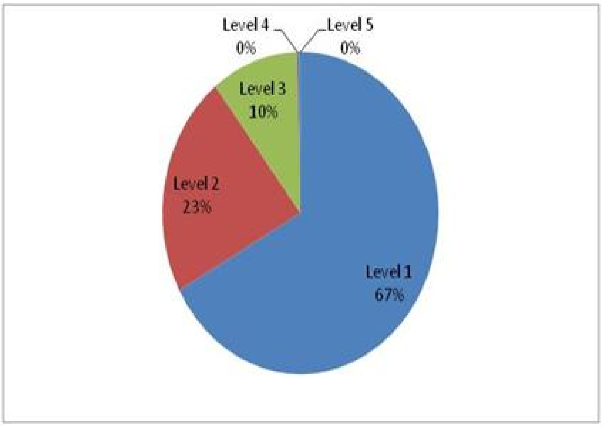


Figure 4. Statistics of Failure Causes of Mail Sorting Machines [Korea Post & J-H Park 2008]

2.2.2 Failu re Steps

We have 5 failure steps [level] o f mail s orting machine o n curr ent m aintenance s ystem such as si mple operation problem & clean of mail sorting machine [MSM] [Step1], replace of simple parts & align & tilt of MSM [Step2], replace of unit and parts & test of MSM [Step3], replace of critical unit and parts of sorting machine & test [Step4], need urgent help of sorting machine manufacture & hot line connection [Step5]. The figure 5 shows the statistics of failure step of all mail sorting machines at the end of 2007. The major failure steps of mail sorting machine based on the figure 5 is step 1 of 67% and step 2 o f 23%. It shows 31 .4 % f ailure rate of stack er par t and 22 .3 % f ailure r ate o f feeder part on mail s orting machine.



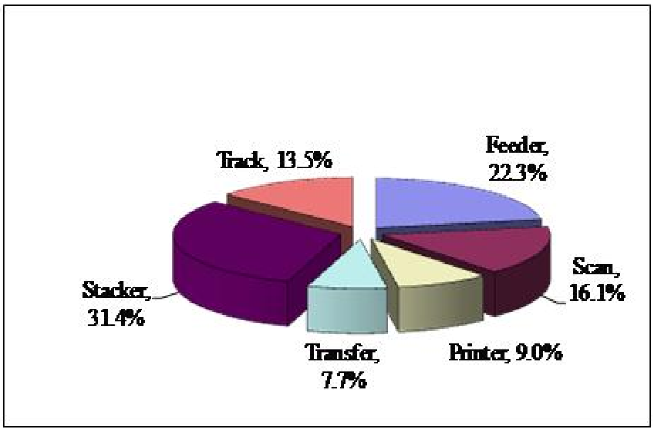


Figure 5. Statistics of Failure Steps [Levels] & Systems of Mail Sorting Machines [Korea Post & J-H Park 2008]

Of course, the failure rate of sorting machine is different from sorting machine such as OVIS, PSM, FSM, and so on. Major failure sorting machine is OVIS/LSM which shows 66.7 % failure rate, and flat sorting machine shows 20% failure rate.

2.3 Maintenance Problem

Even if there are 5 failure steps and 7 failure causes for mail sorting machine, there are no diagnosis and repair manuals of mail sorting machine. Also, there are no standard procedure and guideline of diagnosis and repair of sorting machine. Even there is no part list and replace manual of part for sorting machine. We need stock management and demand and support planning of part for sortin g machine. Other wise, there are many barrier s t o ov ercome to be succ essful in i mproving reliability, and we won't be able to sustain the attention necessary to achieve internal change.

- There is no consistency how parts and units are performing for maintenance.
- Planners dedicated to units do very little routine planning. Instead they are exp editers, on-call supervisors, and when they do plan, it is for outages.
- Maintenance craft skills are deteriorating. No one in the organization is assuring the continuing development of craft skills.
- The remaining central force feels alienated from the unit based maintenance crew.
- The Reliability Engineering Team (usually those who perform the Predictive Maintenance function) are frustrated that their success is limited to those units whose managers understand their value.
- Important measures of planned maintenance and schedule conformance are declining or very stubborn at improving. Operating units have no standard definitions of these measures, and may or may not even measure and record them.
- There is no program and education for professional of maintenance staff, and training centre.
- There is no function and system to share maintenance data which performed diagnosis and repair the sorting machine on each mail distribution centre.
- There is no practical contract for maintenance of sorting machine between operator and manufacture.
- There is no consistency policy, planning and schedule for maintenance of mail sorting machine.

For any of these problems, we consider is based, in general, on centralizing functions that create efficiency and control of in maintenance, and decentralizing functions of work effectiveness in maintenance.

3. MAINTENANCE SYSTEM OF MAIL SORTING MACHINE

We describe to-be co nsiderations for maintenance o f mail so rting machine such as maintenance s ystem, st andard procedure for diagnosis and repair of mail sorting machine, roles and functions of maintenance technical support centre for the training and education of maintenance staff in this section. The following are co nsiderations as principles for efficient maintenance system of mail sorting machine in Korea [6].

- Stability: All mail sorting machines should be operated with reliability and safety maintenance program.

- Professional: Maintenance training program and maintenance technical support should be operated for skill upgrade of maintenance staff, and supported on/off line by maintenance technical support center.
- Standardization: There should be standard diagnosis and repair procedures and manuals, part lists and replace manual, and standard demand and support procedure of part for maintenance of sorting machine.
- Economy: Maintenance system should be cost effective and efficiency.

3.1 Maintenance System

Maintenance system based on maintenance subject of mail sorting machine in Korea classifies 4 types such as maintenance by full internal staff, full outsourcing, internal and outsourcing which has branch office for maintenance of sorting machine on mail distribution center (semi internal type 1), internal and outsourcing which has no branch office for maintenance of mail sorting machine on mail distribution center (semi internal type 2).

In terms of economy, efficiency, reliability & safety of 4 maintenance system types, we assigned weighted value on each maintenance system according to AHP (Analytic Hierarchy Process).

The figure 6 shows this result of maintenance system. The semi internal type 1 has first priority in terms of economy, and full internal staff type has first priority in terms of efficiency. In general we think semi internal type 1 is better for maintenance of mail sorting machine in Korea [J-H Park 2008].

Evaluation Factor	Management System	Weighted Value	Priority
Economy	Internal Staff + (Regular Visiting for Diagnosis) Outsourcing	0.4073	1
	Internal Staff + (Branch Office Work) Outsourcing	0.2315	2
	Full Internal Staff	0.1643	4
	Full Outsourcing	0.1969	3
Efficiency	Internal Staff + (Regular Visiting for Diagnosis) Outsourcing	0.1938	3
	Internal Staff + (Branch Office Work) Outsourcing	0.3050	2
	Full Internal Staff	0.3903	1
	Full Outsourcing	0.1110	4
Reliability & Safety	Internal Staff + (Regular Visiting for Diagnosis) Outsourcing	0.2958	2
	Internal Staff + (Branch Office Work) Outsourcing	0.3542	1
	Full Internal Staff	0.2065	3
	Full Outsourcing	0.1438	4

Management System	Result	Priority
Internal Staff + (Regular Visiting for Diagnosis) Outsourcing	0.357	1
Internal Staff + (Branch Office Work) Outsourcing	0.298	2
Full Internal Staff	0.205	3
Full Outsourcing	0.140	4
	1	

Figure 6. Weighted Value and Priority of Maintenance System Based on AHP [J-H Park & J-H Hong 2008]

We also reviewed cost analysis of 4 types, and the semi internal type 1 is better maintenance system in terms of cost analysis. The figure 7 shows cost analysis and AHP result for maintenance of mail sorting machine in Korea. Cost analysis is based on sorting machine cost, average salary of internal staff and outsourcing member, repair time and cost, and so on.

Management System	Internal Staff Cost [Unit: Million Won]	Outsourcing Cost [Unit: Million Won]	Reliability & Safety Cost [Unit: Million Won]	Total [Unit: Million Won]
Internal Staff + (Regular Visiting for Diagnosis) Outsourcing	6,403	1,656	0	8,059
Internal Staff + (Branch Office Work) Outsourcing	5,502	2,972	-395	8,080
Full Internal Staff	9,073	-	938	10,012
Full Outsourcing	-	10,658	-592	10,656

Maintenance System	Task	Result	Priority
Internal Staff + (Regular Visiting for Diagnosis) Outsourcing	• Self : Daily/Weekly Check/Step 1~3 Repair • Outsourcing : Monthly/Quarterly/Step 4~5 Repair	0.357	1
Internal Staff+ (Branch Office Work, BOW) Outsourcing	• Self : Daily/Weekly Check/Step 1~3 Repair • BOW Outsourcing : Monthly/Quarterly/Step 4~5 Repair	0.298	2
Full Internal Staff	• Self : Step 1-5/Regular Check • Hot Line : High Risk & Critical Repair	0.206	3
Full Outsourcing	• Self : Machine Operation • Outsourcing: Step 1-5/Regular Check	0.140	4
		1.001	

Figure 7. Cost Analysis and AHP Result for Maintenance System of Sorting Machine [J-H Park & J-H Hong 2008]

The figure 8 shows operation and role model of maintenance which is semi internal maintenance type. The failure of step 1 – 3 and daily check is performed by internal staff, the failure of step 4 – 5 and repair of unit replace is performed by outsourcing member (maintenance company), and the critical failure and hot line support for maintenance of mail sorting machine are performed by outsourcing (manufacture).

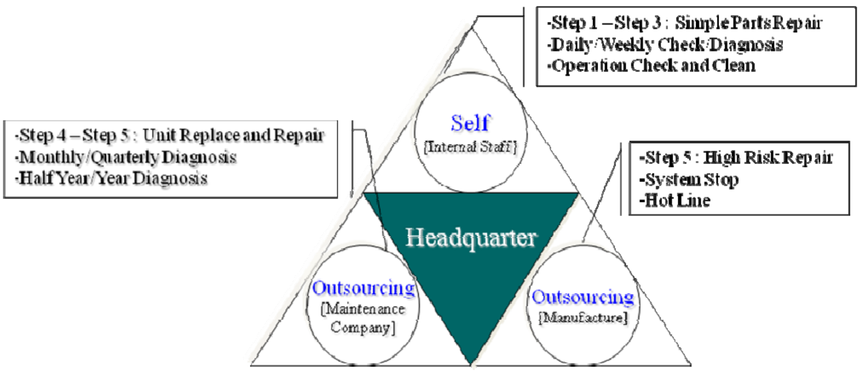


Figure 8. Operation and Role Model of Maintenance [J-H Park 2008]

3.2 Diagnosis and Repair Procedure

It's important to have standard procedure for diagnosis of mail sorting machine on mail distribution centre.

- Diagnosis Procedure

The figure 9 shows standard procedure for diagnosis of mail sorting machine. There are 3 steps for diagnosis of mail sorting machine. The first is the diagnosis preparation step which is the diagnosis planning & schedule, and diagnosis manual preparation of mail sorting machine on mail distribution centre. The second is the diagnosis performing and confirmation step which is the diagnosis manual registration on system, diagnosis performing, and confirmation of mail sorting system. The third is the management step after diagnosis which is the repair of mail sorting machine and update of diagnosis manual on system.

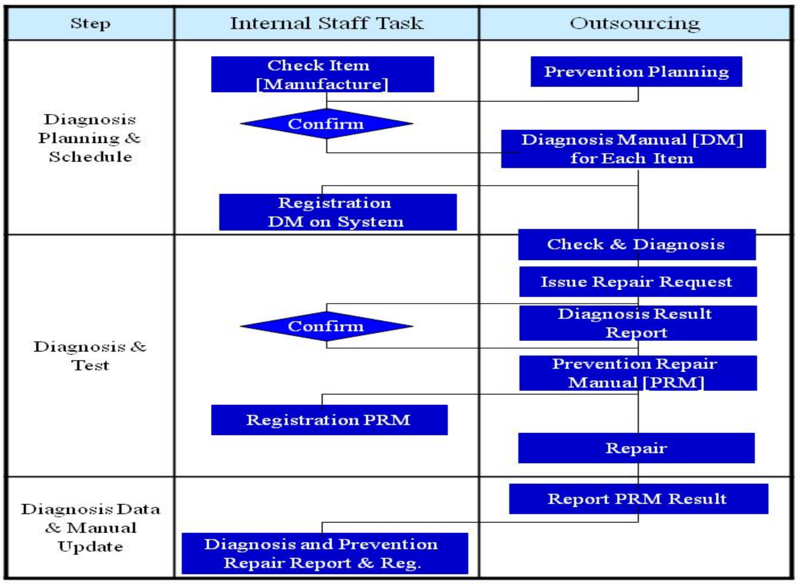


Figure 9. Diagnosis Procedure of Mail Sorting Machine [J-H Park 2008]

- Repair Procedure

It's also important to have standard procedure for repair of mail sorting machine on mail distribution centre. The figure 10 shows standard procedure for repair of mail sorting machine. There are 3 steps for repair of mail sorting machine. The first is the repair request step which is requested the technical support for repair of failure of mail sorting machine by specialist such as out sourcing member in maintenance. In this step it is performed of failure root cause analysis and decision by specialist in maintenance. The second is the repair preparation step which is the repair manual preparation and registration on system, and the preparation of part for repair of mail sorting machine. The third is the repair and confirmation step which is the repair of mail sorting machine and update of repair manual.

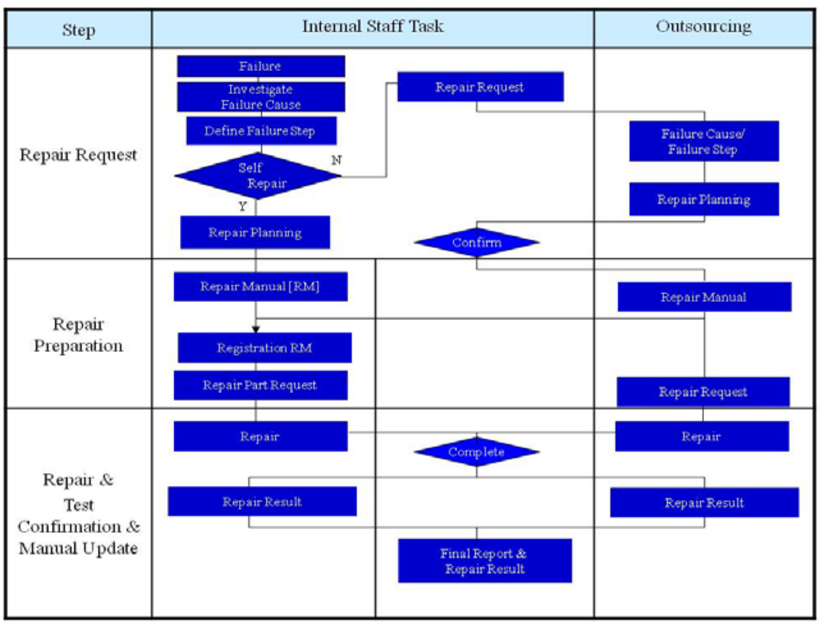


Figure 10. Repair Procedure of Mail Sorting Machine [J-H Park 2008]

3.3 Operation of Maintenance Technical Support Center

A model for organizing maintenance is based on, in general, on centralizing functions that create efficiency and control of work, and decentralizing functions of work effectiveness.

Function of efficiency and control in maintenance includes the following: work prioritization for global resources, work planning, work scheduling for global resources, analysis of work done, preventive and predictive process, maintenance of information t ools, compliance wi th s tandards, central rep orting, skill assur ance. Fun ction of work e ffectiveness in maintenance includes t he following: respon se to immediate nee ds, re cognition and all eviation of equipment, ch ronic problems within the operating context, prioritizing and scheduling area resources.

Otherwise, the training and education of maintenance staff is important to improve the maintenance quality of mail sorting machine. So, maintenance technical support center will support this training program and education on on/off line. There is multi level maintenance training program for maintenance staff, operator, and manager of mail sorting machine. The figure 11 shows the operation and role of maintenance technical support center on semi internal maintenance type 1. The major roles of maintenance tech nical suppo rt center are th e f ollowing: maintenance train ing p rogram a nd edu cation, sto ck management, demand and support of part, on/off line maintenance support & help desk, operation of m aintenance system & environment establishment for share of maintenance data, history management of maintenance data & standardization, standard procedure and manual management for maintenance, and so on.

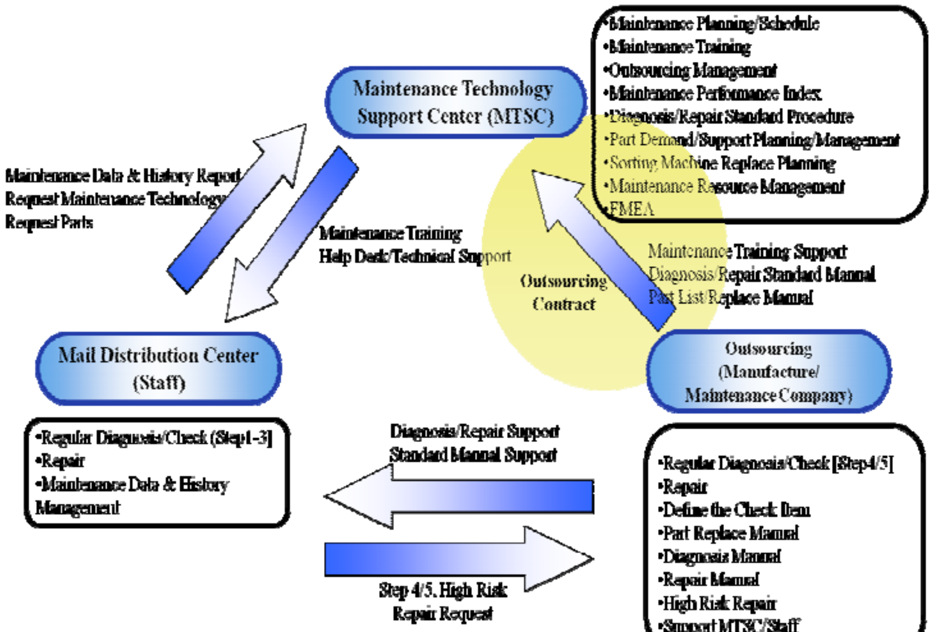


Figure 11. Operation & Role of Maintenance Technical Support Center [J-H Park & J-H Hong 2008]

3.4 Others for Improving of Maintenance

A key point of the efficient maintenance system is stability, professional, standardization, economy. And a key element in maintenance is continuous improvement of the skills of the personnel in maintenance through education and training [1]. There is con tinual need to i mprove th e overall p erformance of equip ment t hrough co mmunication a nd tea mwork. The figure 12 shows the relation among these for improving of maintenance.

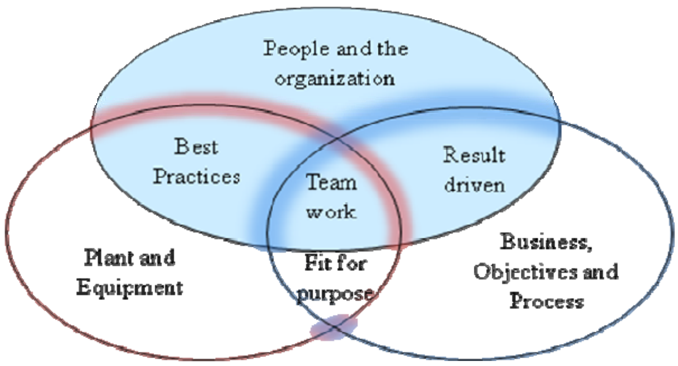


Figure 12. Relation among consideration factors for improving of Maintenance [EFNMS 1999]

Any of these techniques and technologies can be helpful in improving equipment reliability and maintenance.

- P&S (Planning and Scheduling). Planned work is shown to be three times as time-efficient as unplanned work, leading to getting more work done
- PM (Preventive Maintenance). If through proper inspections we can catch and fix problems before failure, reliability should improve
- PdM (Predictive Maintenance). By understanding equipment condition under load conditions we should be able to identify and trend impending failures, replacing components at the end of their useful life ?
- CMMS (Computerized Maintenance Management Systems). Installing a new system should lead to better control and organizational knowledge, giving us the information we need to improve reliability
- TPM (Total Productive Maintenance). Enabling the operator to perform equipment care tasks puts responsibility for problem identification and minor maintenance with the person closest to the equipment
- RCM (Reliability Centered Maintenance). RCM promises to eliminate the source of the equipment defect so that it cannot occur
- RCFA (Root Cause Failure Analysis). When problems occur we can study what went wrong, and stop it from happening again

What any one of these processes in isolation misses is that fundamental behaviors must change for any of them to be systematically applied throughout the maintenance of mail sorting machine on mail distribution center. Some are very complex to learn; some a huge amount of effort to implement. If any step is done with an incomplete understanding, the results will be suboptimal.

4. CONCLUSIONS

The pressure on all types of industry for profitability is rapidly increasing. Management realizes that increased competence at all levels in an organization is necessary to ensure growth. The principal objective of asset management is to optimize the use of capital. Management realizes that increased competence, co-operation and better communication at all levels in an organization leads to more effective maintenance. One of the key factors in continuous improvement is the education and training of all maintenance personnel. To implement best maintenance approaches, it is important to;

- Realise that its development is a combination of applied technical systems along with the competent personnel operating and maintaining them.
- Acknowledge that structured changes are required to reflect the business environment
- Recognize the importance of the involvement of all personnel at all levels for achieving business objectives
- Apply suitable tools to strive for continuously improved equipment effectiveness

- Increase involvement and integration at all levels in order to pursue cost effective maintenance within the framework of life-cycle costing

In this paper, we suggested to-be considerations such maintenance system, and standard procedure of diagnosis and repair for improving of maintenance of mail sorting machine on mail distribution center in Korea. For efficient and economic maintenance system, we reviewed cost analysis of 4 possible maintenance types such as full outsourcing, full internal staff, semi internal type 1, semi internal type 2 based on AHP (Analytic Hierarchy Process). We also suggested the model and role of maintenance technical support center for professional and upgrade of maintenance staff. There are also key functions and others considerations for improving of maintenance.

However, there has been a lot of debate over the years revolving around how maintenance should be measured. The theories vary far and wide, but in essence, most of the debate concerned how to measure the end results of what has been accomplished. Should it be measured on maintenance cost per unit output, cost as a percent of replacement value, or on equipment uptime, etcetera? The debate continues to rage, but we contend that the debate is focused on the wrong timeframe; all of these measurements, as stated, concern themselves with "outcomes" or after the fact measurements.

We also have to consider the development of maintenance operation system for maintenance staff and problem management of mail sorting machine, and definition of maintenance item and cycle based on FMEA (Failure Mode Effect analysis) for further study.

ACKNOWLEDGEMENTS

This work was supported by the Postal Technology R&D program of MKE/IITA. [2006-X-001-02, Development of Real-time Postal Logistics System].

5. REFERENCES

EFNMS Working Group 5. (1999). Future Competence Challenges Related to Maintenance and Asset Management.

S. Bradley Peterson. (2005). The Central Issue: How to Make Distributed Maintenance Work. STRATEGIC ASSET MANAGEMENT INC (SAMI), USA.

S. Bradley Peterson. (2002). Maintenance Outsourcing: A New Model for Operational Excellence. STRATEGIC ASSET MANAGEMENT INC (SAMI), USA.

Ralph D. Hedding. (2005). Measuring Maintenance Effectiveness: The Bulls and Bears. STRATEGIC ASSET MANAGEMENT INC (SAMI), USA.

Korea Post. (2008). Operation and Maintenance. Korea Post Statistical Report, Korea.

J-H. Park et al. (2008). Maintenance Status and TO-BE Considerations of Mail Sorting Machines. ETRI, Korea.

J. H. Hong et al. (2008). Analysis for Efficient Maintenance of Heterogeneous Mail Sorting Machine on Distributed Operation Environment. HanSeong University, Korea2.

SIMILARITY MEASUREMENT FOR LOGISTIC PROCESSES

Bernardo Nugroho Yahya[1] and Hyerim Bae[1*]

[1]Department of Industrial Engineering
Pusan National University
30-san Jangjeon-dong, Geumjong-gu
Pusan, South Korea
email: bernardo@pusan.ac.kr
[*]e-mail: hrbae@pusan.ac.kr (corresponding author)

1. INTRODUCTION

In order to achieve competitive advantage and adapt to dynamic business conditions, each logistics organization can implement customized process to meet customer requirements while complying with their business objectives. Thus, customizability is important characteristic for the organization's competitive advantage. We define customizability as the ability of the logistic party to apply logistics process objectives in many different business conditions. It is a common practice to customize processes from reference processes or templates in order to reduce the time and effort required to design and deploy processes on all levels (Lazovik and Ludwig, 2007). Customizing reference processes usually involves adding, removing or modifying process elements such as activities, control flow and data flow connectors. However, large number of customized process may arise process redundancy. Thus, it encourages organization to carry out process reusability function. Process reusability is the ability to develop a process model once and use it again multiple times within the same process or across multiple processes. Business process modeling enables to reuse of design process and to compose new processes from reused process. Such reusability should be correct, robust and recognizable to users.

Development of a similarity evaluation approach is called for as a precautionary measure limiting the occurrence of redundant processes. Major contribution of this paper is the development of a process similarity measurement for convenient logistic process modeling through elimination of redundant processes. When there is a similar process objective, the model that an operation analysis deems as the most similar structure among existing logistics process can be reused. By using the activity proximity score (APS), it is believed that it can be used to efficiently manage the issues regarding duplicative processes as well as the reuse of any existing logistics process.

2. LITERATURE REVIEW

A business process can be viewed as a collection of decision models, and each process is identified by a type of decision and contains a sequence of processing tasks. The term business process modeling is used to incorporate all activities relating to the transformation of knowledge regarding business systems into models that describe the processes performed by organizations. Process modeling has always been at the core of BPM. Models enable the system to initiate relevant activities so that business objectives can be achieved.

Bae *et al.* (2007) proposed a similarity analysis technique that measures the relationships of adjacent activities among two business processes, named as BPSAT. The similarity, utilizing the process- dependency graph and normalization matrix, represented relationships among activities within a process. However, it expressed no split and merge relationships. Moreover, this technique results in integer values which may cause imprecise analysis when the number of processes is increased. Another process similarity approach proposed workflow clustering based on process similarity (BPClustering) (Jung, Bae and Liu, 2009). Yahya, Bae and Bae (2009) enhanced the idea using a system of proximity score measurement called Business Process Proximity Score Measurement (BP-PSM) (Yahya, Bae and Bae, 2009). Inspired by the sequence alignment concept, which had been introduced ithin the field of protein research and which uses a Point Access Matrix (PAM) and a Block Substitution Matrix (BLOSUM) (Henikoff and Henikoff, 1992; Dayhoff, Schwartz and Orcutt, 1978), Yahya, Bae and Bae (2009) presented a new idea for discovering the best process based on the proximity score measurement (PSM) value. The present paper is an enhancement of research into PSM for process design selection, specifically involving measurement of the activity distance using the transitive distance relationship.

3. SIMILARITY MEASUREMENT USING APS

In this study, we used the business process model defined in Definition 1.

Definition 1. (Process Model)
We define a process p as a tuple of $<A, L>$ and labeling function f, each element of which is defined below.

- $A = \{a_i | i=1,\dots,I\}$ is the set of activities, where a_i is the i-th activity of p and I is the total number of activities in p
- $L \subseteq \{l_{ij} = (a_i,a_j) \mid a_i,a_j \in A\}$ is the set of links, where l_{ij} is the link between two activities a_i and a_j. The element (a_i,a_j) represents the fact that a_i immediately precedes a_j
- a_{i+} is the activity following a_i and a_{i-} is the activity preceding a_i
- For a split activity a_i, such that $|SA_i| > 1$, where $SA_i=\{a_{j+} |(a_i,a_{i+})\in L\}$, $f(a_i)$ = 'AND' if all a_{i+} should be executed; otherwise, $f(a_{i+})$= 'OR'
- For a merge activity a_i, such that $|MA_i| > 1$, where $MA_i=\{a_{j-}|(a_i,a_{i-}) \in L\}$, $f(a_{i-})$ = 'AND' if all a_{i-} should be executed; otherwise, $f(a_{i-})$= 'OR'

Definition 2. (Path, distance, reachability)
Let $PA_{ij} = \{ pa_t|t = 1, 2, 3, \dots, T\}$ where PA_{ij} is a set of paths from an activity a_i to another activity a_j as an edge-connectivity. Since there might be multiple paths between the two activities, an element of the set pa_t is defined as the t-th path with T is the number of paths from a_i to a_j. We denote $pa_t=\{a_i, a_k, a_{k+1}, \dots,a_{k+m},\dots,a_{k+M},a_j\}$, $(a_i, a_k)\in L$, $(a_{k+M}, a_j) \in L$, and for all m, $(a_{k+m}, a_{k+m+1}) \in L$. If there is a path from a_i to a_j, we say that a_i is reachable to a_j, and to represent that reachability, we denote as $a_i \rightarrow a_j$. On the contrary, we use $a_i \parallel a_j$ to represent that a_j is not reachable from a_i. When there is a path pa_t ($\in PA_{ij}$) from a_i to a_j, the distance (d_t) between the two activities equals to $|pa_t|$ -1 or M+1 (Yahya, Bae and Bae (2009)).

We also measure the average distance between two activities that are reachable with multiple paths. Multiple paths exist where there are activities in the split, merge and *iterative* block. Bae et al. (2004) explains more about the block structure as it pertains to parallel activities such as the AND- and OR-splits. Yahya, Bae and Bae (2009) proposed a method to measure the average path distance of AND- and OR- parallel blocks. In the present study, we include *iterative* activity as the part of control flow structure.

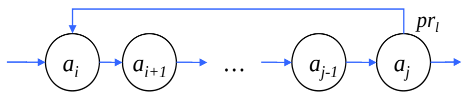


Figure 1. *Iterative*-block abstraction model

Definition 3. (Average Path Distance of *iterative*-block)
Figure 1 shows the iteration pattern by expressing the activity flow from a_j back to a_i. Let assume that the probability of re-execution is pr_l. To measure the activity distance in the *iterative*-block, first we need to calculate the distance between two activities (d_{ij}), which is activity a_i and a_j, then the looping distance. The existence of the *iterative*-block is specified by a process design user at process build-time, and the re-execution probability can be determined from previous data or by expert knowledge. We can calculate the distance between two activities $\overline{d}_{ij}$, both of which are located within the *iterative*-block, by multiplying the distance between a_i and a_j (d_{ij}) with the re-execution probability, as shown in equation (1).

$$\overline{d}_{ij} = d_{ij} + pr_l.d_{ij} + pr_l^2.d_{ij} + \dots = \frac{d_{ij}}{1-pr_l} \quad (1)$$

Definition 4. (Activity Proximity Score: APS)
In this paper, we define q_{ij}^k as the APS between activity a_i and a_j in the k-th process index. The score is obtained by the following equation

$$q_{ij}^k = \frac{h^k(i,j)}{\overline{d_{ij}^k}} \quad \text{where } h^k(i,j) = 1, \text{ if } a_i \rightarrow a_j \text{ in } k\text{-th process, 0 otherwise} \quad (2),$$

$\overline{d_{ij}^k}$ is the distance between activity a_i and a_j in $pa_t \in PA_{ij}$ of the k-th process. Each process has a single value of q_{ij}^k

,k={1,2,3,…,K}, where K is the total number of processes. If there is no relationship between activity a_i and a_j in the k-th process, or it is denoted as $a_i \parallel a_j$, then $q_{ij}^k = 0$.

Definition 5. (Process distance (*D*))

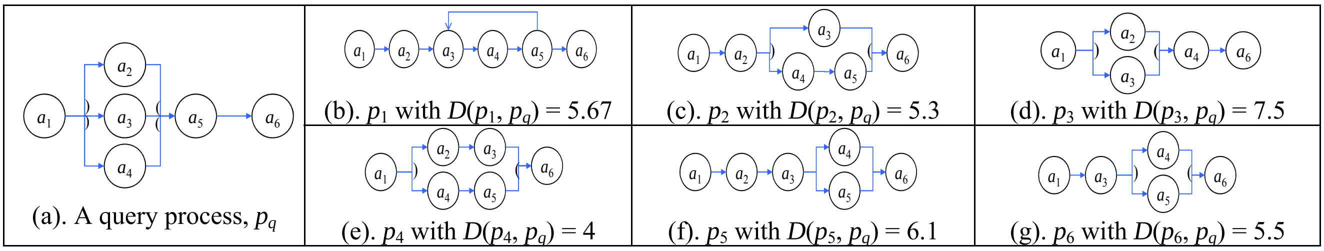
We define process distance (D) between two processes based on the similarity, which can be represented using APS value. Process distance between two processes is defined as the summation of all of the absolute values of the proximity differences of all pairs of activities between processes k and l where $k,l \in K$. It is denoted as follows:

$$D(p_k, p_l) = \sum_{i,j} \left| q_{ij}^k - q_{ij}^l \right| \quad (3).$$

In terms of APS, two processes with smaller value are closer, in other words, we use a smaller-the-better method. Thus, a value of 0 means those two processes are identical in their structure.

4. SIMILARITY MEASUREMENT USING APS

Suppose we want to measure the similarity by finding the most similar process to a query process (p_q) in a repository with K process variants. Process variants are generated from a single process model, each of which fits a certain scenario and context; in other words, the configuration of a particular process variant reflects the specific requirement and circumstance of the process. The process distance (D) can be used to find a similar process, which is the result of the similarity score between p_q and each process among process variants. We presented example of process variants in Figure 1. The process distance (D) value of each process variant compare with query process (p_q) are shown in each cell.



(a). A query process, p_q

(b). p_1 with $D(p_1, p_q) = 5.67$

(c). p_2 with $D(p_2, p_q) = 5.3$

(d). p_3 with $D(p_3, p_q) = 7.5$

(e). p_4 with $D(p_4, p_q) = 4$

(f). p_5 with $D(p_5, p_q) = 6.1$

(g). p_6 with $D(p_6, p_q) = 5.5$

Figure 2. Example of process variants with similarity measurement using process distance

With regard to performance of APS, we are expressing the function of transitive closure proposed by Warshall's algorithm using Big-O notation. We define the total number of activities in a process as N. In order to find the activity pair, the computation complexity becomes $O(N^2)$. To develop efficient solutions to reachability problems, the concept of transitive closure is employed. Transitive closure in graph theory involves the construction of a path, by means of a data structure mechanism, which renders a possible solution to the reachability problem (Sedgewick, 2004). Warshall's algorithm, which proposed a solution of the transitive closure problem, includes the t-th iteration of loop sets in order to find any pairs of activities with indices greater than 1, since there are T paths from a_i to a_j (equation (1)). Thus, there will be $|pa_t|$-1 times (equal to N+1 activities) to measure the path distance from a_i to a_j. The complexity of computing the APS, therefore, is $O(N^3)$.

5. FINAL REMARKS

This paper proposes a method of logistics process similarity measurement. Based on the business structure, we can obtain the APS and compare processes having similar structures. The sequential, parallel and iterative structures were studied in this proposed method. In the application, we can gain some of the advantages by using APS with regard to process reusability, business alignment, business process risk management, business process outsourcing and business partner selection.

There are some issues for further research. Comparison analysis with other methods is necessary to show the merit of this approach. Further semantic analysis is a promising and contributory approach in the logistics research domain. In this approach, each relationship score can connect to the ontology of any semantics to generate the required process.

ACKNOWLEDGEMENTS

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (The Regional Research Universities Program/Research Center for Logistics Information Technology).

6. REFERENCES

Bae, H. and Seo, Y. (2007) 'BPM-based integration of supply chain process modeling, executing, and monitoring', *International Journal of Production Research*, vol. 45, No. 11, pp. 2545 – 2566.

Bae, J., Liu, L., Caverlee, J., Zhang, L.J. and Bae, H. (2007) 'Development of Distance Measures for Process Mining, Discovery and Integration', *International Journal of Web Services Research*, vol. 4, issue 4, pp. 1-17.

Bae, J., Bae, H., Kang, S. and Kim, Y. (2004) 'Automatic Control of Workflow Processes Using ECA Rules', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 16, No. 8, pp. 1010-1023.

Buckley, F. and Harary, F. (1990) *Distance in Graphs*, Addison-Wesley Publishing Company.

Chopra, S. and Meindl, P., (2006) *Supply Chain Management*, 3rd edition, Prentice Hall: Upper Sadle River, NJ

Dayhoff, M.O., Schwartz, R.M. and Orcutt, B. (1978) 'A model of evolutionary change in proteins', *Atlas of Protein Sequence and Structure, vol. 5, suppl. 345-352.* National Biomedical Research Foundation, Silver Spring, MD.

Henikoff, S. and Henikoff, J.G. (1992) 'Amino acid substitution matrices from protein blocks' *Proc Natl Acad Sci U.S.A 89*, 10915-9

Jung, J.Y., Bae, J. and Liu, L. (2009) 'Hierarchical Clustering of Business Process Models', *International Journal of Innovative Computing, Information and Control*, Vol. 5, No. 12, pp. 4501-4511.

Lambert, D.M., Cooper, M.C. and Pagh, J.D. (1998), 'Supply Chain Management: Implementation Issues and Research Opportunities', *International Journal of Logistics Management*, Vol. 9, No. 2. pp. 1-19.

Lazovik, A. and Ludwig, H. (2007) 'Managing Process Customizability and Customization: Model, Language and Process', *in LNCS 4831, WISE (2007)*, pp. 373-384.

SCC (2009) '*Supply-Chain Operation Reference Model: Overview of SCOR version 9.0*'. Supply Chain Council, Inc.

Sedgewick, R. (2004) *Algorithm in Java*, 3rd edition, Pearson Education Inc.

Smith, H. and Fingar, P. (2003) *Business Process Management, the third wave*, Morgan-Kiffer Press.

Van der Aalst, W.M.P, Weijters, T. and Maruster, L. (2004) 'Workflow Mining: Discovering Process Models from Event Logs', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 16, No. 9, September 2004..

Yahya, B. N., Bae, H. and Bae., J. (2009) 'Process Design Selection using Proximity Score Measurement', *3rd International Workshop on Collaborative Business Processes (CBP '09).* Germany.

Session E3: Seaport and Transportation 5

·Day2: Sep. 16 (Thu.)

·Time: 09:00 - 10:20

·Chair: Frank Meisel

·Room: Iris, 4F

LOGMS

The 1st International Conference on Logistics and Maritime Systems

September 15th–17th, 2010 Busan, Korea

Lin

BENCHMARK GENERATION FOR QUAY CRANE SCHEDULING PROBLEMS

Frank Meisel[1]* and Christian Bierwirth[1]

[1] Martin-Luther-University
Germany
e-mail: frank.meisel@wiwi.uni-halle.de
christian.bierwirth@wiwi.uni-halle

Abstract: Scheduling of quay cranes at container terminals is a field of growing interest in research and practice. In the literature, we find diverse models for quay crane scheduling that consider the service of container vessels at different aggregation levels as well as a multitude of different solution procedures. Currently, however, there exists no basis to compare the quality of planning which is achieved by the different models. Furthermore, there is no benchmark platform available that allows comparing different solution procedures and investigating conditions where a procedure performs well or poorly. In this talk, we describe a structured process for generating a broad range of test instances with application to diverse quay crane scheduling models. A Java implementation of the scheme is used to generate new benchmark problems, which enable evaluations of QCSP models and algorithms in diverse directions.

COMPARING VARIOUS STRATEGIES FOR TRANSPORTER POOLING IN CONTAINER TERMINALS

Vu Duc Nguyen and Kap Hwan Kim*

Department of Industrial Engineering
Pusan National University
30 Jangjeon-dong, Geumjeong-gu
Busan 609-735, South Korea
e-mail: duc.ise@pusan.ac.kr and kapkim@pusan.ac.kr
* Corresponding author

Abstract: In container terminals, containers are transported between vessels and storage blocks by transporters. To improve the utilization of transporters as well as the operation efficiency of container terminals, the pools of transporters to transfer containers are constructed and transporters in a pool are utilized for carrying out a pre-specified group of tasks. This study compares various strategies for constructing the pools: pools for a single quay crane, pools for a group of quay cranes assigned to a vessel, pools for a group of quay cranes for multiple adjacent vessels, a big pool for all quay cranes. Besides, the different heuristic rules to assign pools to storage blocks are suggested to consider travel distance of transporters at quay side as well as that in the yard. Also, the opportunities of dual cycle operations are considered in these heuristic rules. A number of scenarios are generated and compared with each other by a simulation study in terms of the total delay times of QCs, the total travel times, and the total empty travel times of transporters.

1. INTRODUCTION

In container terminals, there are three main types of handling equipments involved in terminal operations: quay cranes (QCs), transporters (yard trucks - YTs), and yard cranes (YCs). For an unloading operation, a container is unloaded from a ship by a QC and transported to assigned block by a transporter. In the yard, it is picked up from and stacked into a storage block by a YC. For a loading operation, a container is handled in the reverse direction to the unloading operation. During the unloading and loading operations, containers are transported between vessels and storage blocks by transporters. To improve the utilization of transporters as well as the operation efficiency of container terminals, the pools of transporters are constructed and transporters in a pool are utilized for carrying out a pre-specified group of tasks. This study compares various strategies for constructing the pools: pools for a single QC, pools for a group of QCs assigned to a vessel, pools for a group of QCs for multiple adjacent vessels, a big pool for all the QCs. In pools for a single QC, every YT is assigned to a single dedicated QC. In this case, the control system is simple, but the utilization of YTs may decrease because of the long empty travels. However, in the other pools, a YT can be assigned to more than one QC. During a cycle of a YT, after delivering a container from a QC to a storage block, the YT can deliver another from a YC to another QC. Thus, the control system may be more complicated and the higher utilization of YTs is expected. Besides, a heuristic algorithm to assign QCs to pools is suggested by considering the travel distance of YTs at the quay side as well as that in the yard. Also, the opportunities of dual cycle operations are considered in the heuristic algorithm. In dual cycle operations, a loading operation immediately performs after an unloading operation or vice versa so that the empty travel of handling equipments is minimized. Recently, the dual cycle operations have been applied for improving the terminal operation in container terminals. Zhang and Kim (2009) studied the sequencing problem not only for stacks under a hatch cover but also for hatches. They attempted to minimize the number of operation cycles of a QC for discharging and loading containers in a ship-bay by maximizing the number of dual cycle operations.

For the vehicle dispatching problem, Bae and Kim (2000) proposed a pooled dispatching strategy for automated guided vehicles (AGVs) in port container terminals. A mathematical formulation is developed and a heuristic algorithm is suggested to obtain a near optimal solution within a reasonable amount of computational time. Also, the single cycle and the dual cycle operations in both the seaside and landside operations are analyzed. Grunow et al. (2006) introduced a simulation study of AGV dispatching strategies in a seaport container terminal, where AGVs can be used in either single-carrier or dual-carrier mode. They compared a typical, on-line dispatching strategy adopted from flexible manufacturing systems with a pattern-based, off-line heuristic algorithm. Briskorn et al. (2006) presented an alternative formulation of the AGV assignment problem that does not include due-times and is based on a rough analogy to inventory management; they proposed an exact algorithm for solving the formulation. Nguyen and Kim (2009)

proposed a mathematical formulation of the dispatching problem for automated lifting vehicles (ALVs). A heuristic algorithm was suggested and the solutions of the heuristic algorithm were compared with optimal solutions. Also, the numerical experiments were performed to analyse and evaluate the performance of heuristic algorithm.
Simulations have been used as a powerful tool for analyzing the performance of port container terminals in complex dynamic environments. Various levels of detail and the uncertainties in container terminals can be expressed in simulation models. Much research on the development of simulation models of container terminals has been published (Cho, 1985; Yun and Cho, 1999; Tahar and Hussain, 2000). Hartmann (2004) introduced an approach for generating scenarios that can be used as input data for simulation models of port container terminals.

This paper is organized as follows. Section 2 introduces the ship operation and suggests various types of pools of YTs. A number of scenarios are generated and compared by a simulation study in Section 3. Section 4 discusses some conclusions and issues for further researches.

2. THE SHIP OPERATION AND POOLING TRANSPORTERS BY A HEURISTIC ALGORITHM

2.1 Ship Operation

Before a ship arrives at a port container terminal, all information on inbound and outbound containers is sent to the terminal by a shipping agent. Then, based on the information, a sequence list of the unloading and loading operations of individual containers is constructed. When the ship actually arrives at the terminal, ship operations are usually carried out based on the loading and discharging sequence list.

At the quay side, an unloading operation by a QC starts with a release of a container onto a YT, while a loading operation by a QC starts with a pickup of a container from a YT. At the yard side, a loading operation by a YC starts with a release of a container onto a YT, while an unloading operation by a YC starts with a pickup of a container from a YT. Containers are transferred between quay side and yard side by YTs. Besides, the dual cycle operations can be performed to enhance the terminal operation productivity. Figure 1 shows various types of operations performed by YTs. In the single operations as shown in Figure 1a, when a YT delivers a loading operation or unloading operation independently so that an empty movement is needed at every cycle. While, in dual cycle operations, a YT performs an unloading operation after finishing a loading operation. That means an unloading operation and a loading operation are combined in one cycle, thus the empty movements may be reduced in comparison with the single operation. In Figures 1b, 1c, and 1d, there are empty movements of YTs between two QCs, two blocks, between two QCs and two blocks, respectively. Figure 1e illustrates the full dual cycle without empty movements between QCs and blocks, which means a QC performs loading and discharging operations in dual cycles and the storage locations of the containers to be discharged are the same as those for the containers to be loaded. During the dual cycle operations, the "savings" in the empty travel time can be calculated by deducting empty travel movements between two QCs, two blocks, or between QCs and blocks (in Figure 1b, 1c, and 1d, respectively) from that between a QC and a block in Figure 1a. In case of Figure 1e, there are no empty travel movements between QCs and blocks, and thus the "savings" is the entire empty movements between a QC and a block.

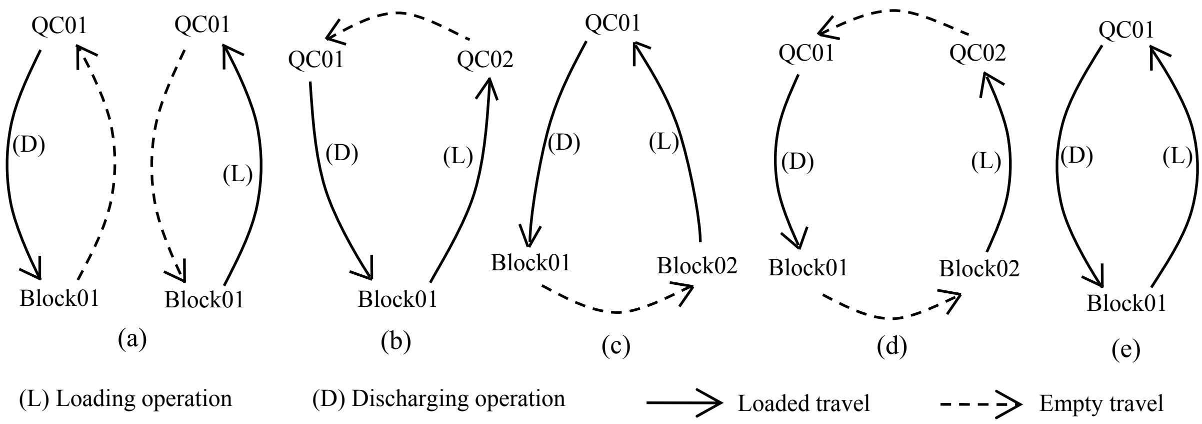
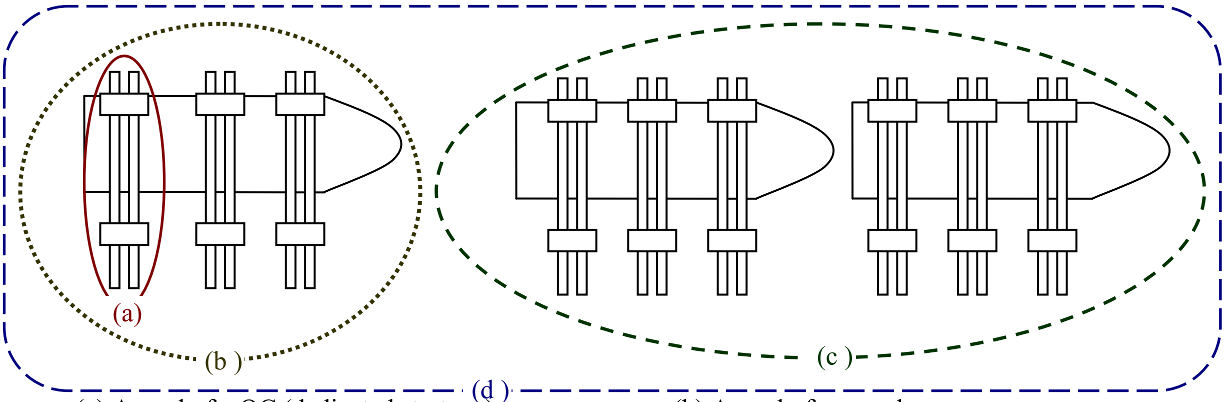


Figure 1. Different types of operations performed by YT.

2.2 Pooling YTs by a Heuristic Algorithm

This study assumes that a fixed number of transporters are assigned to each QC. However, by grouping QCs, transporters assigned to QCs in each group are shared by all the QCs in the same group. The transporters shared by QCs in each group are said to be in the same pool. Figure 2 illustrates four strategies for constructing the pools of transporters: pools for a single QC, pools for a group of QCs assigned to the same vessel, pools for a group of QCs for multiple adjacent vessels, a big pool for all the QCs.



(a) A pool of a QC (dedicated strategy) (b) A pool of a vessel

(c) A pool of two adjacent vessels (d) A pool of all vessels (full pooled strategy)

Figure 2. Definition of traditional pools.

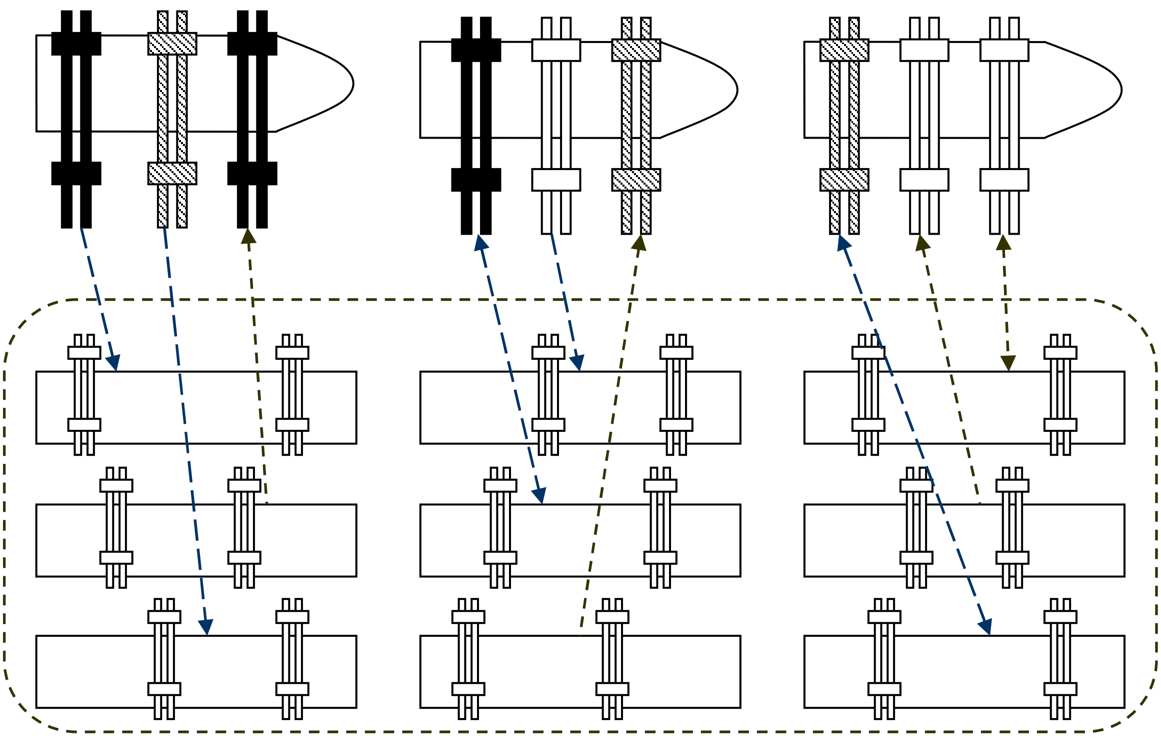


Figure 3. Flexible pooling strategy.

Besides, a flexible pooling strategy suggested in this study is illustrated in Figure 3. Figure 3 illustrates an example of three pools of transporters. Each pool consists of three QCs. QCs with the same pattern inside represent QCs whose tasks are served by the same pool of transporters. An arrow from a QC to a block represents that the transporters assigned to the QC should move discharged containers to the corresponding block. Arrows from blocks to QCs represent movements of outbound containers. Arrows with heads at both ends represent dual cycle operations of QCs. Note that the pooling strategy in Figure 3 considers the travel distance of transporters at the quay side as well as that in the yard. Also, the opportunities of dual cycle operations are considered. In this paper, the main idea of the pooling algorithm is as follows. This algorithm will be in charge of grouping QCs into multiple pools. Pooling algorithm will be triggered whenever a QC starts a task for a new group of containers. Containers which are to be moved between the same QC and block are said to be in the same group. The algorithm may start from estimating the effect of pooling by calculating the "savings" in its empty travel when a QC is inserted in a pool. The QCs will be grouped in a way of maximizing the "savings" resulting from the pooling.

Assumptions

1. The sequence of discharging and loading containers for each QC is predetermined and cannot be changed.
2. The storage locations for discharging containers are predetermined.

Notation

p_i : pool index, $i = 1, \ldots, I$.

Q_i : set of QCs in pool i, $i = 1, \ldots, I$.

n_{ij}^U : number of unloading containers of QC j, $j = 1, \ldots, J$ in pool i, which are discharged simultaneously with operations for the new group of containers of the current QC.

n_{ij}^L : number of loading containers of QC j, $j = 1, \ldots, J$ in pool i, which are loaded simultaneously with operations for the new group of containers of the current QC.

n_k : number of containers in the new group of the current QC k; where, $n_k \geq n_{ij}^U + n_{ij}^L, \forall i = 1, \ldots, I,\ j = 1, \ldots, J$.

d_j : expected travel distance between QC j and the storage location in the yard of containers for QC j, which are transferred simultaneously with containers of the current QC.

d_{ij}^Q : expected travel distance between QC i and QC j.

d_{ij}^Y : expected travel distance between the storage locations of containers for QC i and j in the yard.

s_i: "savings" of pool i when a QC is inserted into pool $i = 1, \ldots, I$.

Procedure for the Heuristic Algorithm

Step 0: Initializing

Construct initial pools for transporters each of which consists of all the transporters assigned to QCs deployed to each vessel (pools for a group of quay cranes assigned to a vessel).

Step 1: Estimating the "savings", s_i, $i = 1, \ldots, I$.

When a QC k of pool i starts loading or discharging a new group of containers and $|Q_i| > 1$, the "savings" for insertion of QC pool i, $i = 1, \ldots, I$ is calculated as follows.

Let $m_i = \sum_{j \in Q_i} n_{ij}^U - \sum_{j \in Q_i} n_{ij}^L$.

If the new group of containers of QC k is for unloading operation,

And if $m_i \geq 0$ then $s_i = 0$.

Otherwise, s_i is estimated as follows.

- $s_i = 0$;
- $j = 1$;
- $n = \min(-m_i, n_k)$;
- Repeat

  If $j \in Q_i$ and $n > 0$ then

  If $n_{ij}^L < n$, then

  $s_i = s_i + n_{ij}^L.(d_j + d_k - d_{jk}^Y)$;

  $n = n - n_{ij}^L$;

  Else

  $s_i = s_i + n.(d_j + d_k - d_{jk}^Y)$;

  $n = 0$;

  End;

  End;

  $j = j + 1$;

  Until $j > J$ or $n = 0$.

If the new group of containers of QC k is loading operation,

And if $m_i < 0$ then $s_i = 0$.

Otherwise, s_i is estimated as follows.

- $s_i = 0$;
- $j = 1$;
- $n = \min(m_i, n_k)$;
- Repeat

  If $j \in Q_i$ and $n > 0$ then

  If $n_{ij}^U < n$ then

  $s_i = s_i + n_{ij}^U.(d_j + d_k - d_{jk}^Q)$;

  $n = n - n_{ij}^U$;

  Else

  $s_i = s_i + n.(d_j + d_k - d_{jk}^Q)$;

  $n = 0$;

  End;

  End;

  $j = j + 1$;

  Until $j > J$ or $n = 0$.

Step 2: Selecting the maximum "savings", s_i, $i = 1, \ldots, I$.

Select the pool with maximum "savings" to which QC k is to be inserted.

3. SIMULATION EXPERIMENTS

The operation of the container terminal was modeled in a simulation model in detail. When a ship arrives at the terminal, it is assigned to a berth if there is an available berth. Otherwise, the ship must wait until a berth becomes available. When the ship enters a berth, a pre-specified number of QCs will be assigned to the ship. Then, a loading and unloading sequence will be generated for each QC. With reference to the specified sequence, QCs can start to unload and load containers.

The operation scenarios modelled in the simulation program are described below. At the quay side, QCs will unload or load containers only when YTs are available under QCs. When a YT arrives at a designated QC, it must wait for either pickup or drop-off of a container by a QC. At the yard side, YTs will first arrive at transfer points (TPs) in the side of the block. The YT will wait at a TP for a YC to pickup an inbound container or release an outbound container. When a YT finishes transporting a container, it will return to the parking area to await the next assignment.

For the loading operation, a YC picks up a loading container from a specified slot in the yard. For the unloading operation, when a YC receives an inbound container from a vehicle, it will stack the container in a slot in the yard. The loading and unloading operation will continue until all containers are transferred between the ship and the storage yard. When all operations are completed, the ship will leave the berth.

In the following, the simulation model is introduced. The wharf has 3 berths and 9 QCs. The yard consists of 9 storage blocks to each of which two YCs of the same size are deployed. Each block has TPs on which YTs can wait for transferring containers. The travel times of YTs between pickup and drop-off points are assumed to be constant, and the congestion of YTs on guide paths is not considered. The total number of containers transferred by QCs is about 1350 during a run of the simulation.

The four scenarios of pooling transporters, as shown in Figure 2, are constructed in the simulation experiment: scenario A – pools for a single QC, scenario B – pools for a group of QCs assigned to a vessel, scenario C – pools for a group of QCs for two adjacent vessels, scenario D – a big pool for all the QCs. Also, scenario E – pooling transporters by the suggested heuristic algorithm shown in Figure 3 was conducted in these experiments. The data considered in simulation experiments which the unloading and loading operations are completely mixed between stacks in a bay to make dual cycle operations of QC. The performance measures considered in simulation experiments are the total delay times of QCs, the total travel time of YTs, and the total empty travel time of YTs.

Table 1. Comparisons between scenarios for different pooling strategies.

Scenarios	Total delay times of QCs (in sec.)	Total travel time of YTs (in sec.)	Total empty travel time of YTs (in sec.)
A 13	1,910	374,473	155,600
B 12	5,728	**360,050**	145,083
C 12	1,531	361,943	145,385
D	**120,638**	363,669 14	4,643
E 12	1,091	361,848	**143,501**

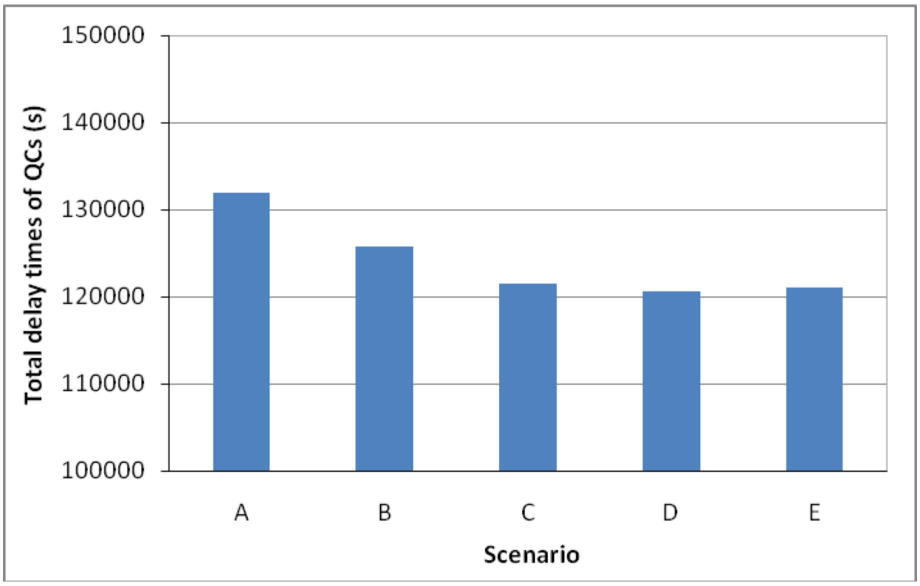


Figure 4. Comparison of total delay times of QCs among five scenarios.

The results from simulation experiments shown in Table 1 indicated that the dedicated pooling was the worst. When the scope of the pool became wider, the performances were better. Figure 4 compares the total delay time of QC operations among different scenarios. When the number of QCs assigned to a pool increased (like scenario C, D and E), the total delay time of QCs reduced. The average reductions of delay times in scenario E were 8.2% and 3.7% compared with scenario A and B, respectively. It means that the improvement of scenario E compared with scenario B which maintained the initial pools of scenario E to the end was about 3.7%. Figures 5 and 6 compare the total travel times and the total empty travel times of YTs for various scenarios. The pooling strategies outperformed the dedicated strategy.

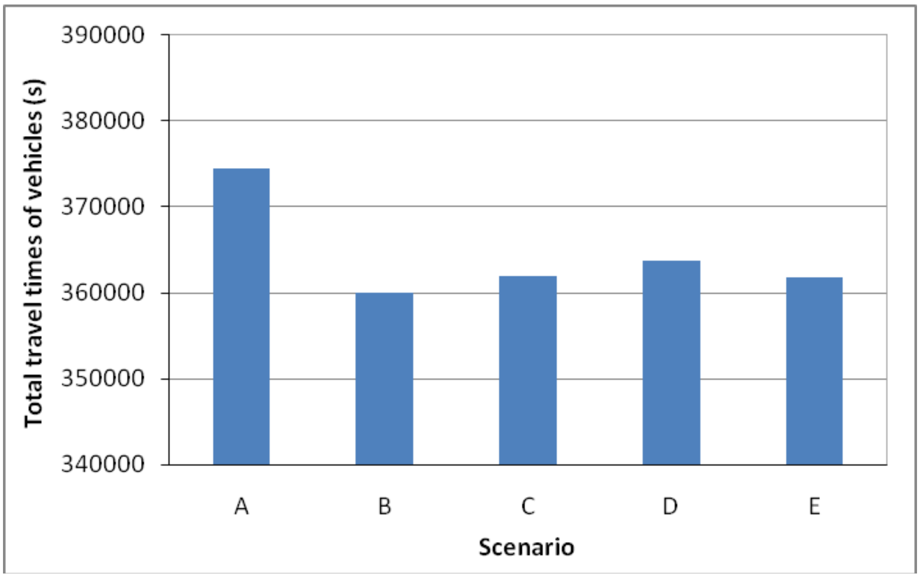


Figure 5. Comparison of total travel times of YTs among scenarios.

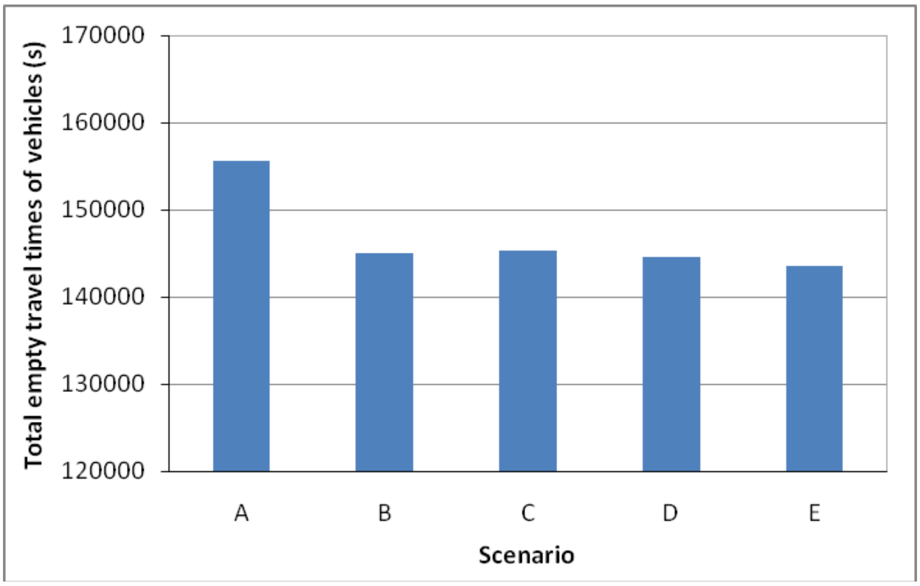


Figure 6. Comparison of total empty travel times among scenarios.

4. CONCLUSIONS

This study compared various strategies for constructing pools of transporters in container terminals: pools for a single QC, pools for a group of QCs assigned to a vessel, pools for a group of QCs for multiple adjacent vessels, a big pool for all the QCs. Besides, a heuristic algorithm to assign QCs to pools is suggested to consider travel distance of YTs at quay side as well as that in the yard. Also, the opportunities of dual cycle operations are considered in the heuristic algorithm.

A simulation study was conducted to compare and evaluate the performances of various strategies. From the simulation results, it was found that the proposed algorithm outperforms the dedicated pooling strategy and other pooling strategies except the big pool strategy in the total delay times of QCs and shows a performance similar to that of the big pool. In terms of the total travel times of YTs and the total empty travel times of YTs, the proposed algorithm also showed one of the best performances among tested strategies. This study mainly focuses on the operation of transporters; however, the integrated operations of vehicles, YCs, and QCs may be addressed in further research.

ACKNOWLEDGEMENTS

This study was supported by the Korean Ministry of Education & Human Resources Development through the Research Center of Logistics Information Technologies (LIT).

5. REFERENCES

Bae J. W. and Kim K. H. (2000). A Pooled Dispatching Strategy for Automated Guided Vehicles in Automated Port Container Terminals. *International Journal of Management Science*, 6(2), 47-67.

Briskorn D., Drexl A. and Hartmann S. (2006). Inventory-Based Dispatching of Automated Guided Vehicles on Container Terminals. *OR Spectrum*, 28, 611-630.

Cho, D. W. (1985). A Computer Simulation Model for Container Terminal Systems. *Journal of the Korean Institute of Industrial Engineers*, 11, 173-187.

Hartmann, S. (2004). Generating Scenarios for Simulation and Optimization of Container Logistics. *OR Spectrum*, 26, 171-192.

Grunow M., Günther H. O., and Lehmann M. (2006). Strategies for Dispatching AGVs at Automated Seaport Container Terminals. *OR Spectrum*, 28, 587-610.

Nguyen V. D. and Kim K. H. (2009). A Dispatching Method for Automated Lifting Vehicles in Automated Port Container Terminals. *Computers & Industrial Engineering*, 56(3), 1002-1020.

Tahar, R. M. and Hussain, K. (2000). Simulation and Analysis for the Kelang Container Terminal Operations. *Logistics Information Management*, 13, 14-20.

Yun, W. Y. and Choi, Y. S. (1999). A Simulation Model for Container-Terminal Operation Analysis Using an Object-oriented Approach. *International Journal of Production Economics*, 59, 221-230.

Zhang H.-P. and Kim K. H. (2009). Maximizing the Number of Dual-Cycle Operations of Quay Cranes in Container Terminals. *Computers & Industrial Engineering*, 56(3), 979-992.

STORAGE BLOCK ASSIGNMENT FOR IMPORT CONTAINERS IN AUTOMATED SEAPORT CONTAINER TERMINALS

Rico Gujjula*, Hans-Otto Günther

Department of Production Management
TU Berlin – H 95
Straße des 17. Juni 135
Berlin, Germany
e-mail: {rico.gujjula, hans-otto.guenther}@tu-berlin.de

1. INTRODUCTION

The CTA in Hamburg and the ECT in Rotterdam are highly automated seaport container terminals which deploy automated guided vehicles (AGVs) to transport containers from quayside to yard and vice versa. The so-called quayside transport is one of the key operations in a terminal. A rather neglected problem in literature with a considerable impact on the performance of the quayside transport is storage block assignment for import containers (cf. Gujjula & Günther, 2008). In this paper we discuss an algorithm to solve this assignment problem in order to increase the performance of quayside transport. Solutions of the proposed algorithm can be used to benchmark results which were generated by use of priority rules.

2. PROBLEM DESCRIPTION

In this paper we consider an automated seaport container terminal which comprises, AGVs, automated quay cranes and storage blocks which are placed perpendicularly to the quay. Each storage block deploys two automated yard cranes, which are exclusively serving sea- or landside, respectively. Thus, in this paper we consider only the crane serving the seaside.

There are three types of container movements within the terminal. Import containers are moved from quayside to landside, export containers from landside to quayside and transshipment containers from quayside to the storage blocks for intermediate storage and back to the quayside.

As for the import direction, a vessel arrives with the containers at the allocated berthing place and quay cranes are assigned for discharging. A quay crane picks up a container by use of a spreader from the vessel, lifts it up with a hoist, moves the container to the landside using the trolley and puts the container onto an AGV. Afterwards, the container is transported to a storage block at the stacking area and is picked up by the yard crane directly from the AGV. The yard crane puts the container into a position at the stack until it will be requested for landside transport.

As for export containers, a container arrives by truck or train and is stacked at a storage block until it will be requested by a vessel at the quayside. Then the container is picked up from the stacking area, is transported to the quay crane that charges the vessel and is put onto a slot on the vessel. The quay crane where the container is charged and the slot on the vessel are determined by the stowage plan of the vessel which was generated in advance.

Quayside transport is concerned with the transport of containers from quayside to yard and vice versa and covers several logistics problems like the dispatching of AGVs, routing and traffic control. Optimization methods and heuristics for quayside transport problems focus on the enhancement of quay crane productivity and the reduction of vehicle travel time (cf. Steenken *et al.*, 2004). A matter of importance is to synchronize cranes and vehicles to avoid waiting time of terminal equipment.

One of the many problems with an impact on the quayside transport performance is the storage block assignment problem for import containers. As most data is inaccurate or in-complete upon arrival of such containers, no stacking strategies like for export containers are applied, hence every storage block with free capacity is acceptable for any import container (Steenken *et al.*, 2004). When an import container was loaded onto an AGV a storage block needs to be assigned at the latest. Clearly, the choice of a storage block affects other terminal operations, first and foremost, AGV dispatching and yard crane dispatching. For the remainder of this paper we assume that the applied procedures for AGV and yard crane dispatching are given. The proposed heuristic makes use of this input in order to anticipate decisions by these dispatching procedures.

3. PROPOSED ALGORITHM

Whenever a storage block is needed for an import container, each storage blocks has a list of pending jobs for either import or export containers. The execution order of these jobs is determined by the yard crane dispatching procedure.

The algorithm evaluates for each storage block the inclusion of a new job for this container. At first, the arrival time of its AGV at the block will be determined. Afterwards, yard crane dispatching is anticipated using the set of known jobs and the new job for the import container in order to determine the start and finishing time of each job.

There are three options how the algorithm can choose a storage block for the import container. With the first option, the algorithm chooses the closest storage block with the smallest increase in the makespan if the import job is executed at this block. With the second option, the behavior of AGVs leaving a block with an export container is first determined using the anticipated job order for this block. From this order one can extrapolate at which time the AGVs arrive at the target quay crane and, thus, the waiting time for these quay cranes can be derived. The algorithm chooses the storage block so that the overall waiting time for all quay cranes is minimal. With the third option, the algorithm additionally anticipates AGV dispatching for empty AGVs which dropped an import container at this block. In this case, quay crane waiting times can be derived and the algorithm chooses a storage block such that the overall quay crane waiting time is minimal. Note that the algorithm needs to assume that all processing and travel times are given in advance.

In the presentation for this paper we give an overview of the current state of this algorithm and present preliminary results for the performance of the proposed algorithm using the three possible options and use these results to benchmark the performance of priority rules.

4. REFERENCES

Gujjula, R. and Günther, H.O. (2008). The impact of storage block assignment for import containers on AGV dispatching in highly automated seaport container terminals. *Proceedings of the 2008 IEEE International Conference on Industrial Engineering and Engineering Management*, IEEE, 1739-1743.

Steenken, D., Voß, S., Stahlbock, R. (2004). Container terminal operations and operations research – a classification and literature review. *OR Spectrum, 26:* 3-49.

Session E4: Industrial Session 1

· **Day2: Sep. 16 (Thu.)**

· **Time: 09:00 - 10:30**

· **Chair: Kwang Ryel Ryu**

· **Room: Grand Ballroom, 5F**

PLANNING AND OPERATING AUTOMATED CONTAINER TERMINALS: LESSONS LEARNED

Sonke Hartmann [1]*

[1] Hamburg School of Business Administration
Germany
e-mail: soenke.hartmann@hsba.de

Abstract: Several container terminals with a substantial degree of automation have started operation during the last decade or are currently under construction. Popular terminal designs include the combination of automated stacking cranes with either automated guided vehicles (AGVs) or straddle carriers. In addition to the container handling equipment, also the concepts for terminal operating systems (TOS) have become more and more advanced. Facing the increasing complexity, terminals typically employ simulation as a tool to cope with the dynamics of the container handling processes.

In this presentation, simulation and optimization approaches for container terminals are discussed. It is based on experience gained during several consulting and research projects for automated container terminals. We discuss scheduling concepts for both automated equipment and human resources and deal with modeling issues, data availability, and algorithms. Moreover, we consider simulation models and their applications to container terminals. Thereby, we identify some pitfalls and limitations of simulation in practice. Finally, the role of research in container terminal logistics and its contribution to practice is discussed.

Session F1: SCM 5

·**Day2: Sep. 16 (Thu.)**

·**Time: 10:40 - 12:00**

·**Chair: Jirapan Liangrokapart**

·**Room: Camellia, 5F**

COLLABORATION BASED RECONFIGURATION OF PACKAGE SERVICE NETWORK WITH MULTIPLE CONSOLIDATION TERMINALS

Chang Seong Ko[1], Eun Mi Yoon[1], Ki Ho Chung[2], and Hyun Jeung Ko[3*]

[1]Department of Industrial & Management Engineering
Kyungsung University
Daeyeon-dong, Nam-gu
Busan, 608-736
Korea
e-mail: csko@ks.ac.kr, eunmi1023@naver.com

[2]Department of Management Information Systems
Kyungsung University
Daeyeon-dong, Nam-gu
Busan, 608-736
Korea
e-mail: khchung@ks.ac.kr

[3]Department of Logistics
Kunsan National University
1170 University Avenue
Gunsan, Jeonbuk, 573-701
Korea
e-mail: hjko@kunsan.ac.kr

Abstract: The market competition of package deliveries in Korea is severe because a large number of companies have entered into the market. However, a package delivery company in Korea generally owns and operates a number of service centers and consolidation terminals for high level customer service although some service centers as cost centers cannot create profits due to low volume. This challenge can be overcome by collaboration strategy based on win-win principles. Therefore, this study suggests an approach to the reconfiguration of a package service network with respect to collaboration strategy, reassignments of service centers to consolidation terminals, and adjustments of their cutoff times. For this we propose an integer programming model and a genetic algorithm based solution procedure for participated companies to maximize their profit. An illustrative numerical example in Korea is presented to demonstrate the practicality and efficiency of the proposed model.

1. INTRODUCTION

The collaboration or partnership is becoming a popular competitive strategy to be adopted in all business sectors. Some of well known examples can be seen in the air transportation system such as Sky Team, Star Alliance, and Oneworld as well as in sea transportation such as CKYH-The Green Alliance, Grand Alliance, and so on. In addition, the supply chain management regards the concept of collaboration as a critical factor for its successful implementation. In particular, the market of express package deliveries in Korea has been rapidly expanded according to the progress of TV home shopping and internet buying and selling. Accordingly, various sized domestic express companies have been established, and various foreign companies also have entered into the Korean express market. As a result of the surplus of express companies, they are struggling with remaining competitive at a reasonable price with appropriate level of customer satisfaction. In addition, some of big companies are strongly carried out low service price policy by the favor of scale economics.

In Korea, an express company generally operates its own service network which consists of customer zones, service centers, and consolidation terminals. Customer zones refer to geographical districts in which customers either ship or receive packages and are typically covered by a service center. And a service center receives customer shipment requests and picks up parcels from customer zones and then the packages are waited until its cutoff time for transshipment in bulk to a consolidation terminal. In this way, the service center acts as a temporary storage facility

* Corresponding author

connecting customers to a consolidation terminal. At the consolidation terminal, the packages delivered from a number of service centers are combined, screened, sorted, and then loaded onto delivery trucks for their destinations so that it acts as a transshipment hub. Thus, the productivity and service level of an express package service are highly related to how to efficiently manage service centers and consolidation terminals while maintaining appropriate level of customer satisfaction.

A number of express companies, however, are operating under-utilized service centers in certain customer zones for achieving high level service even if they cannot create profits due to low volume as cost centers. This challenge can be overcome by the use of strategic partnership with respect to win-win principles. The strategic partnership works in such a way that express companies operating under-utilized service centers in some places agree to operate only single service center in a specific customer zone, but the decision is made by considering the loss-gain tradeoff between closing and keeping service centers of each company while maximizing the overall operating efficiency of express service networks and fairly allocating throughput volumes of companies involved in partnership. In addition to the partnership, the efficiency of a service center mainly depends on flexibility in setting the cutoff time for the customer's last-minute shipping request. An unnecessarily early cutoff time forces some last-minute customers to give up the next day delivery service and consequently creates inconvenience to customers. On the other hand, an unreasonably late cutoff time leaves service centers little room for handling, sorting, and arranging transportation to the consolidation terminal, and it is likely to delay the succeeding delivery processes.

Therefore, this study suggests an approach to the reconfiguration of express service network with respect to strategy partnership of merging service centers among companies involved, multiple-consolidation terminals for reassigning service centers to consolidation terminals in each company, and adjustments of their cutoff times. For this we propose an integer programming model and a genetic algorithm based solution procedure for allowing them to maximize their incremental profit. An illustrative numerical example with reduced data sets from an express company in Korea is presented to demonstrate the practicality and efficiency of the proposed model.

2. RETERATURE REVIEW

There are a number of research areas related to the design of service network for express package delivery. One of them can be the issue of freight consolidation which has been investigated in numerous analytical studies. Freight consolidation refers to a transportation option that combines a number of frequent, small shipments destined for a similar geographical region into a single large shipment in order to reduce shipping cost per unit and capitalize on freight-rate discount programs. Consolidation can be cost-effective especially when substantial freight-rate differentials exist between small and large shipments. To exploit consolidation opportunities, Powell and Sheffi (1983) initiated a load planning model linking a number of consolidation terminals. However, they did not consider links between terminals and customers. Similarly, Powell (1986) formulated the load planning problem for less-than-truckload (LTL) motor carriers as a fixed charge network design problem. On the other hand, Schneider et al. (1972) were among the first to determine the minimum cost location of urban consolidation terminals. To extend this study, Min (1994) considered both international and multiple objective aspects of a consolidation terminal location-allocation problem. These two studies, however, focused on the issues of location and direct shipment without considering the possibility of serving multiple customers on a single delivery tour.

While the mentioned consolidation studies were primarily concerned with the spatial links among terminals and customers, another aspect of research topics addressed the issue of determining the length of holding time at the consolidation terminal for ordered items. For example, Jackson (1981) used a simulation model to establish the maximum number of days for which orders can be held to take advantage of consolidation. Other pioneering works include Cooper (1983), Buffa (1987), and Higginson and Bookbinder (1994) who dealt with the question of how long customer orders should be held and/or what quantities should be accumulated before consolidated loads are shipped or released.

Despite numerous merits, a few consolidation studies consider spatial network design problems linking terminals to customers and determining the cutoff time for order aggregation simultaneously. In the meantime, Leung et al. (1990) presented a mathematical model and its solution procedure for solving point-to-point delivery problems. However, their study did not take into account door-to-door service coverage problems with cutoff time restrictions on consolidation. The study conducted by Cheung et al. (2001) was the first to examine a service network design problem encountering express couriers such as DHL Hong Kong. They proposed a hybrid optimization/simulation model that aimed to maximize service coverage and service reliability by adjusting cutoff time. This study thus considered spatial network design problems linking terminals to customers and determining the cutoff time for order aggregation simultaneously. Ko et al. (2010) developed an integer programming model and a genetic algorithm to determine the cutoff time at each service center according to the spatial proximity of service centers to customers and the assignment of service centers to a consolidation terminal. To extend this study, Lee at al. (2010) considered the aspects of extensions of terminal capacities. These two studies, however, focused on the issues of only the cutoff time adjustments without considering the possibility of strategic partnership with respect to win-win principles.

To overcome the shortcoming of prior studies, the current study considers the problem of reconfiguring an express package service network with the merger of under-utilized service centers, multiple-consolidation terminals for reassigning service centers to consolidation terminals, and adjusting cutoff time of service centers to maximize the profit.

3. MODEL DESIGN

This section describes a mixed-integer programming model which maximizes the expected profit of an express service networks based on strategic partnership. The basic concept of strategic partnership is to run a single service center at a low volume zone by compromise among involved companies while considering the loss-gain tradeoff. Figure 1 describes the graphical description of the restructured service network. In addition, the cutoff time in this study is time of day before which a customer' order is assured for a next day delivery. The extension of cutoff time for express service centers can provide the company with increase of total sales, but it may also decrease dissatisfying customer needs due to work delay in the consolidation terminal.

Figure 1. A Restructured Service Network based on partnership

Typically, customers' requests of service centers vary according to time of day and the type of their locations such as commercial areas, industrial areas, and residential areas. Demand of express services usually peaks around the daily cutoff time and is high in commercial and industrial areas compared to residential area. Since the demand is concentrated around cutoff times, service centers located in high volume areas want to expand their cut-off time. Underlying assumptions of the proposed model are as follows:

- The operating time of a consolidation terminal is fixed from 8:00 p.m. to 24:00 p.m.
- Once cutoff time is extended, customer orders increase with a constant rate; on the other hand, as cutoff time is set early, express requests decline with a constant rate.
- The current cutoff time of service center is set at 6:00 p.m.; an incremental increase or decrease in cutoff time is 30 minutes; the cutoff time can be shortened to 1.5 hours and extended up to 2 hours.
- The travel time between the service center and the consolidation terminal is constant at any time of a day.
- There is one and only one truck that has a large enough capacity to accommodate shipment loads from a service center to a consolidation terminal.
- The processing time of ordered items at a service center or a consolidation terminal is proportional to the total volume of ordered items.
- The shipments are processed according to the FIFO rule at the consolidation terminal.
- All the outgoing packages from a service center are shipped to the assigned consolidation terminal.
- The customers' orders of closed service centers are assigned to a nearest opened facility.

4. SOLUTION PROCEDURE

Considering that the proposed model in this study belongs to a class of NP-complete problems, we develop a heuristic solution procedure based on genetic algorithm (GA). GA is known as a stochastic solution search procedure which is proven to be useful for providing a good solution for solving combinatorial problems using the concept of evolutionary computation imitating the natural selection and biological reproduction of animal species (Goldberg, 1989; Gen and Cheng, 2000). The subsections will elaborate on the development of GA for the proposed model.

4.1 Encoding and Genetic Operators

The design of a suitable chromosome representation of a candidate solution is the first step for GA application since it dictates probabilistic transition rules for each chromosome to create a population of chromosomes. Each chromosome developed in this study is based on single dimensional array that consists of binary values, representing decision variables associated with the merger of under-utilized service centers, reassignment of service centers to consolidation terminals, and the incremental change in cut off time at service centers. For example, the representation of a chromosome is illustrated in Figure 1. This chromosome describes that there are 60 service centers, out of which 30 are operated by Company A and the rest 30 belongs to Company B. This means that there are two express companies and each company provides express service in 30 customer zones. In addition, we assumed that each company currently operates its service network with 10 underutilized service centers so that they are the candidate facilities for collaboration with respect to win-win principle without hurting its profit structure while meeting customers' service level. Each company also runs multiple consolidation terminals.

In the chromosome, the closing or keeping decisions as merger with respect to partnership and reassignments of service centers to consolidation terminals are made depending on the value of single gene (0 or 1). The adjustment of cutoff times is determined by the combination of three genes. The three genes indicate eight possible values from 0 to 7 so that there are eight possible incremental changes in cutoff time ranging from a maximum reduction of an hour and half to an extension of up to two hours. For example, a set of 0-0-0 refers to 90 minute reduction, 0-0-1 for 1 hour reduction, 1-1-1 for 2 hour extension, and so on.

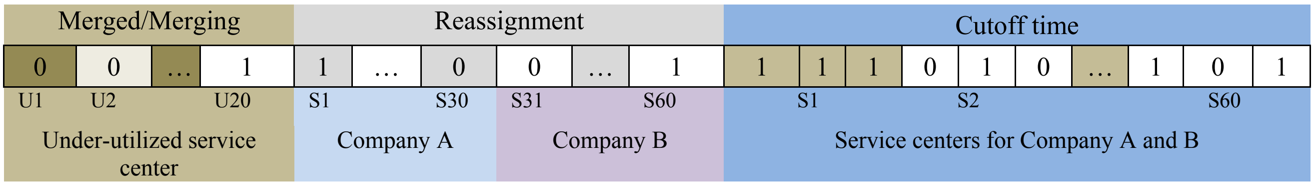


Figure 2. A Chromosome Representation

Four genetic operators are used in the proposed GA such as cloning, parent selection, crossover, and mutation operators. The cloning operator copies 20 percent of the current best chromosomes to a new population. And a binary tournament selection method for a parent selection is used, which begins by forming the two teams of chromosomes. Each team consists of two chromosomes randomly drawn from the current population. The two-best chromosomes that are taken from one of the two teams are chosen for crossover operations. As such, two offspring was generated and entered into a new population. For crossover, the three-point crossover is used in which one point is used for an under-utilized service center, the second for reassignment, and the last one for the cutoff time. The three locations of the crossover points were randomly selected from any genes of a chromosome and then swapped segments of the two parents' strings to produce two children. The mutation operator first randomly selects a bit value of any gene on a chromosome, and then, flips a bit value from 0 to 1, or from 1 to 0 to achieve a good level of diversity in each generation.

4.2 Fitness Function

The decoded chromosome generates a candidate solution and its objective value based on the fitness function. The fitness value is a measure of the goodness of a solution with respect to the original objective function and the extent of infeasibility by adding a penalty to the original objective function. To elaborate, the original objective function is comprised of the goals of maximizing expected incremental profit of overall service network for all companies involved by balancing companies' profit structure.

In particular, the proposed GA first generate possible closing or keeping under-utilized service centers, reassignment of service center to consolidation terminals, and cutoff time extensions. Then, based on a set of decision variables, a fitness value of chromosome were obtained by applying solution procedure. The detailed overall procedure is described as follows:

Step 1: Given the partnership based merging decisions, the shipment requests of a closed service center are assigned to the nearest opened facility, the arrival time of each delivery truck at consolidation terminal is calculated. Based the these settings, sorting operations in accordance with the arrival times of delivery trucks originated from the service centers at the consolidation terminal using the FIFO rule are conducted.

Step 2: Given the sequence of shipments in Step 1, recalculate the task starting and finishing time of each shipment by reversing orders according to the capacity of a consolidation terminal. As such, the starting and completing time of whole process at each consolidation terminal is made. Afterward, if the re-calculated total time of delivery and documentation is infeasible, the cutoff time is adjusted to make the completion time of a current job smaller than the starting time of an immediately succeeding job. This sub-procedure is used for providing improvement of quality of a candidate solution.

Step 3: If the total incremental cutoff time surpasses the total processing time of the consolidation terminal, a penalty is added to the original objective function so that this chromosome is identified as one of infeasible solutions.

5. MODEL EXPERIMENTS

5.1 An Example Problem

An example problem was solved for model application for which its data set is summarized in Table 1 and Table 2. There were two companies such as Company A and Company B. Each company had 30 service centers with two consolidation terminals. Each consolidation terminal processed the shipments from 15 service centers based on geographical proximity. In addition, a company was running 10 underutilized service centers as cost centers.

Currently, the cutoff time was set uniformly at 6:00 p.m. for all the service centers. These service centers were classified into three categories with respect to their shipping volume of express services. For instance, service centers located in residential areas were categorized as “1” where customer demand for express courier service is relatively low; service centers located in industrial areas were categorized as “2” where demand is at the medium level; service centers located in high-demand commercial areas were categorized as “3”. A consolidation terminal of each company started operating at 6:00 p.m. for the shipments of trucks based on first come-first served rule.

Company A currently processes the total of 40,000 units per the day in the consolidation terminals from 8:00 p.m. to 12:00 p.m., and Company B processes the total 39,100 units per the day from 8:00 p.m. to 12:00 p.m. Under this setting, the problem was solved by using the proposed GA that sets parameter values through a series of experiments. These parameters are that the number of population size is varied from 300 to 500; maximum number of generations is from 150 to 300 according to the size of population; the rate of cloning is 20%; the mutation rate varies from 1% to 5% as the number of a generation increases.

Table 1. The current operation of Company A (unit: hour)

	(1)	(2)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Current Operation				
												(11)	(12)	(13)	(14)	(15)
	i	X	Y	fi	ai	α	PCT	SC_CA	SC_PC	Di	TR	T_AT	T_SL	T_ST	T_PC	T_FT
Terminal-1	9	125	16	1900	3	0	18	3800	0.5	179	2.98	21.48	0.00	21.48	0.19	21.67
	10	117	75	1000	2	0	18	2000	0.5	112	1.87	20.37	1.31	21.67	0.10	21.77
	11	145	122	900	3	99.6	18	1800	0.5	93	1.55	20.05	1.72	21.77	0.09	21.86
	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
	26	63	271	550	1	51.8	18	1100	0.5	138	2.30	20.80	2.46	23.26	0.06	23.32
	27	129	275	1550	1	0	18	3100	0.5	154	2.57	21.07	2.25	23.32	0.16	23.47
	28	165	269	950	1	0	18	1900	0.5	184	3.07	21.57	1.91	23.47	0.10	23.57
Terminal- 2	1	215	85	350	2	91.8	18	700	0.5	100	1.67	20.17	0.00	20.17	0.04	20.20
	2	221	126	700	2	97	18	1400	0.5	53	0.88	19.38	0.82	20.20	0.07	20.27
	3	264	115	800	2	89.6	18	1600	0.5	49	0.82	19.32	0.96	20.27	0.08	20.35
	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
	16	228	225	1250	2	0	18	2500	0.5	97	1.62	20.12	2.49	22.61	0.13	22.73
	17	275	235	850	2	67.3	18	1700	0.5	110	1.83	20.33	2.40	22.73	0.09	22.82
	29	192	258	2300	3	0	18	4600	0.5	166	2.77	21.27	1.55	22.82	0.23	23.05
	30	209	287	2500	3	0	18	5000	0.5	178	2.97	21.47	1.58	23.05	0.25	23.30

Columns descriptions : (1) = service center; (2) = coordinates of the location; (3) = current demand; (4) = 1(residential), 2(industrial), 3(commercial); (5) = fixed cost; (6) = closing time of the service center; (7) = hourly capacity, (8) = processing time in service center; (9) = distances between service center and the consolidation terminal; (10) = travel time between service center and the consolidation terminal; (11) = arrival time of a courier truck at the consolidation terminal; (12) = waiting time of a courier truck at the consolidation terminal; (13) = task starting time of a courier truck at the consolidation terminal; (14) = processing time in terminal; (15) = task completion time of a courier truck at the consolidation terminal.

Table 2. The current operation of Company B (unit: hour)

	(1)	(2)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	Current Operation				
												(11)	(12)	(13)	(14)	(15)
	i	X	Y	fi	ai	α	PCT	SC_CA	SC_PC	Di	TR	T_AT	T_SL	T_ST	T_PC	T_FT
Terminal-1	42	170	162	650	1	66	18	1300	0.5	68	1.13	19.63	0.37	20.00	0.07	20.07
	43	204	188	800	3	74.1	18	1600	0.5	76	1.27	19.77	0.30	20.07	0.08	20.15
	44	272	174	1000	2	0	18	2000	0.5	158	2.63	21.13	0.00	21.13	0.10	21.23
	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
	58	165	269	1100	1	0	18	2200	0.5	74	1.23	19.73	2.86	22.59	0.11	22.70
	59	192	258	1900	3	0	18	3800	0.5	90	1.50	20.00	2.70	22.70	0.19	22.89
	60	209	287	2300	3	0	18	4600	0.5	136	2.27	20.77	2.13	22.89	0.23	23.12
Terminal-2	31	215	85	400	2	91.8	18	800	0.5	160	2.67	21.17	0.00	21.17	0.04	21.21
	32	221	126	650	2	97	18	1300	0.5	125	2.08	20.58	0.62	21.21	0.07	21.27
	33	264	115	850	2	89.6	18	1700	0.5	179	2.98	21.48	0.00	21.48	0.09	21.57
	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
	51	87	105	1300	2	0	18	2600	0.5	38	0.63	19.13	4.40	23.53	0.13	23.66
	53	25	75	500	1	88.4	18	1000	0.5	130	2.17	20.67	2.99	23.66	0.05	23.71
	54	24	182	2100	3	0	18	4200	0.5	128	2.13	20.63	3.08	23.71	0.21	23.92

Columns descriptions : (1) = service center; (2) = coordinates of the location; (3) = current demand; (4) = 1(residential), 2(industrial), 3(commercial); (5) = fixed cost; (6) = closing time of the service center; (7) = hourly capacity, (8) = processing time in service center; (9) = distances between service center and the consolidation terminal; (10) = travel time between service center and the consolidation terminal; (11) = arrival time of a courier truck at the consolidation terminal; (12) = waiting time of a courier truck at the consolidation terminal; (13) = task starting time of a courier truck at the consolidation terminal; (14) = processing time in terminal; (15) = task completion time of a courier truck at the consolidation terminal.

5.2 Test Results

For the Company A, the results showed that total six under-utilization service centers were merged such as S1, S2, S11, S12, S13, and S19 reflecting $511.3 savings by reducing operating costs of service centers and the total throughput was achieved up to 43,122.5 units. And given the $1 profit for each unit, this generated the total profit of $43,633.8 resulting in the increase of 9.1%.

In addition to the savings by merger, the effectiveness of running both consolidation terminals was largely improved which means that the starting and completing times of Terminal-1 are 19:45:06 and 23:57:06 compared to the times of current system of 21:28:08 and 23:34:02; for the Terminal-2, the times are 19:54:00 and 23:58:08 compared to 20:10:02 and 23:18:00. By adjusting the cutoff times of service centers according to the characteristics of customer zone, the average driver-waiting times at consolidation terminals was also reduced, which are from 2.10 hour to 0.04 hour for the Terminal-1 and from 1.35 hour to 0.05 for the Terminal-2. Table 3 showed the detailed schedule for service centers and consolidation terminals.

Table 3. The GA test results for Company A after applied strategic partnership (unit: hour)

Terminal-1						Terminal-2					
i	CT	LT	T_AT	T_ST	T_FT	i	CT	LT	T_AT	T_ST	T_FT
27	16.50	16.93	19.49	19.49	19.76	15	17.00	17.45	19.62	19.62	19.90
26	16.50	17.47	19.77	19.77	19.98	14	18.50	19.04	19.80	19.90	20.10
21	18.50	19.04	20.00	20.00	20.22	16	18.00	18.50	20.12	20.12	20.37
20	18.50	19.04	20.14	20.22	20.81	17	17.50	18.37	20.20	20.37	20.66
18	19.00	19.60	20.80	20.81	21.11	4	18.50	19.04	20.70	20.70	20.94
24	19.00	19.60	21.07	21.11	21.66	3	19.00	20.19	21.00	21.00	21.38
23	18.00	18.88	21.38	21.66	21.89	8	19.00	19.60	21.27	21.38	21.74
25	19.00	19.58	21.78	21.89	22.40	6	18.00	18.50	21.78	21.78	22.28
10	19.50	20.11	21.98	22.40	22.64	7	19.00	19.60	22.05	22.28	22.55
22	20.00	20.65	22.65	22.65	23.24	5	19.00	19.55	22.35	22.55	22.78
9	19.50	20.15	23.13	23.24	23.73	29	19.00	19.60	22.37	22.78	23.33
28	20.00	20.60	23.67	23.73	23.96	30	19.50	20.15	23.12	23.33	23.98

Columns' descriptions: (1) = service center; (2) = closing time of the service center; (3) = leaving time of a truck from the service center (4) = arrival time of a truck at the consolidation terminal, (5) = task starting time of a truck at the consolidation terminal, (6) = task completion time of a truck at the consolidation terminal

For the Company B, the results showed that total four under-utilization service centers were merged such as S33, S47, S53, and S56 reflecting $322.5 savings by reducing operating costs of service centers and the total throughput was 45,641.3 units. And the solution generated the total profit of $45,963.8 resulting in the increase of 9.1%.

In addition to the savings by merger, the effectiveness of running both consolidation terminals was substantially improved which said that the starting and completing times of Terminal-1 are 19:08:04 and 24:00:00 compared to the times of current system of 20:00:00 and 23:07:02; for the Terminal-2, the times are 18:50:04 and 23:58:02 compared to 21:10:02 and 23:55:02. By adjusting the cutoff times of service centers, the average driver-waiting times at consolidation terminals was also reduced, which are from 1.66 hour to 0.14 hour for the Terminal-1 and from 2.05 hour to 0.11 for the Terminal-2. Table 4 showed the detailed schedule for service centers and consolidation terminals.

Table 4. The GA test results for Company B after applied strategic partnership (unit: hour)

Terminal-1						Terminal-2					
i	**CT**	**LT**	**T_AT**	**T_ST**	**T_FT**	**i**	**CT**	**LT**	**T_AT**	**T_ST**	**T_FT**
57	17.00	17.45	18.88	18.88	19.14	51	17.50	17.96	18.60	18.60	18.84
52	18.50	19.04	19.20	19.20	19.66	41	18.00	19.10	19.98	19.98	20.31
58	17.50	17.98	19.21	19.66	19.86	38	18.00	18.50	20.10	20.31	20.65
45	18.50	19.03	19.86	19.86	20.20	40	18.50	19.04	20.24	20.65	20.91
50	19.00	19.58	20.31	20.31	20.88	35	16.50	16.93	20.56	20.91	21.10
42	18.50	19.63	20.76	20.88	21.18	49	18.50	19.83	20.90	21.10	21.42
55	18.00	18.50	20.87	21.18	21.58	32	18.00	19.04	21.12	21.42	21.69
46	19.00	19.58	21.13	21.58	21.86	36	18.00	18.50	21.68	21.69	22.17
48	19.50	20.15	21.52	21.86	22.20	54	19.00	19.60	21.73	22.17	22.67
59	20.00	20.70	22.20	22.20	22.73	39	19.50	20.15	22.47	22.67	23.11
60	19.50	20.15	22.42	22.73	23.33	37	20.00	20.70	23.15	23.15	23.49
43	20.00	21.27	22.54	23.33	23.74	31	19.50	20.65	23.32	23.49	23.67
44	20.00	20.65	23.28	23.74	24.00	34	19.00	19.58	23.41	23.67	23.97

Columns' descriptions: (1) = service center; (2) = closing time of the service center; (3) = leaving time of a truck from the service center (4) = arrival time of a truck at the consolidation terminal, (5) = task starting time of a truck at the consolidation terminal, (6) = task completion time of a truck at the consolidation terminal

To sum up, the proposed model showed the usefulness of collaboration strategy based on the win-win principle without hurting the profit structure. Table 5 showed the summary of the test results by applying the proposed model.

Table 5. The summary of test results (unit: hour)

		Company A		Company B	
		Terminal-1	Terminal-2	Terminal-1	Terminal-2
Current	Service centers	15	15	15	15
	Start	21.48	20.17	20.00	21.17
	Finish	23.57	23.30	23.12	23.92
	Throughput	20,850 units	19,150 units	21,350 units	17,750 units
	Waiting time	2.10	1.35	1.66	2.05
	Total profit	$40,000		$39,100	
After	Service centers	12	12	13	13
	Start	19.76	19.90	19.14	18.84
	Finish	23.96	23.98	24.00	23.97
	Throughput	22,111.3 units	21,011.3 units	24,695 units	20,947 units
	Waiting time	0.04	0.05	0.14	0.11
	Savings	$511.3		$322.5	
	Total profit	$43,633.8		$45,963.8	

6. CONCLUSIONS

The collaboration or partnership is becoming a popular competitive strategy to be adopted in all business sectors. Some of well known examples can be seen in the air transportation system such as Sky Team, Star Alliance and Oneworld as well as in sea transportation such as CKYH-The Green Alliance, Grand Alliance, and so on. In addition, the supply chain management regards the concept of collaboration as a critical factor for its successful implementation. In this regard, this paper showed the applicability of collaboration in the design of express service network.

In Particular, express delivery market in Korea is currently facing the low margin competition since a large number of domestic and foreign express companies have been set up and some of big companies are strongly carried out low service price policy by the favor of scale economics. As a result, small-medium companies need to figure out a new management strategy toward the trends, which is able to make their businesses more efficient and competitive. To cope with this difficulty, the partnership or collaboration seems to be useful tool for the companies to reduce operating costs as much as possible by minimizing cost centers in their express service network.

In addition to the collaboration, small-medium express companies need to capture the last-minute customers who are willing to pay price for delivery services by optimize the cutoff times of service centers. The cut-off time is time of day before which a customer' order is assured for a next day delivery. Extension of cut-off time for express service centers can provide the company with increase of total sales, but it can also decrease dissatisfying customer needs due to work delay in the consolidation terminal.

Therefore, this paper develops a mathematical model and its efficient solution procedure based on a genetic algorithm. It aims to maximize the expected profit from express package services by the strategy partnership of merging service centers, the reassignments of service centers to consolidation terminals, and adjustments of their cutoff times. The proposed model considers not only merging under-utilized service centers with respect to win-win principle but also determining the exact length of holding time for collection at a service center and the detailed working schedule at the consolidation terminal. The numerical example showed that the partnership based express service network design approach could provide express companies involved a number of benefits, resulting in the increase of sales, the reduction of waiting time for the truck drivers at the consolidation terminal, and the reduction of operating costs for remaining price competitiveness. As a further study, the authors are planning to conduct capacity variations of multiple consolidation terminals.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2009-0089781).

5. REFERENCES

Powell, W.B., and Y. Sheffi(1983). The load planning problem of motor carriers: problem description and a proposed solution approach. *Transportation Research A,17(6):*471-480.

Powell, W.B.(1986). A local improvement heuristic for the design of less-than-truckload motor carrier networks. *Transportation Science, 20(4):*246-257.

Schneider, J.B., Symons, J.G., and M. Goldman(1972). Planning transportation terminal systems in urban regions. *Transportation Research, 6*:257-273.

Min, H.(1994). Location analysis of international consolidation terminals using the analytic hierarchy process. *Journal of Business Logistics, 15(2):*25-44.

Jackson, G.C.(1981), Evaluating order consolidation strategies using simulation, *Journal of Business Logistics, 2(2):*110-138.

Cooper, M.C.(1983), Freight consolidation and warehouse location strategies in physical distribution systems, *Journal of Business Logistics, 4(2)*:53-74.

Buffa, F.P.(1987), Transit time and cost factors: their effects on inbound consolidation, *Transportation Journal*, 27:50-63.

Higginson, J.K. and J. H. Bookbinder(1994), Policy recommendations for a shipment consolidation program", *Journal of Business Logistics*, *15(1)*:87-112.

Leung, L., Magnanti, T., and V. Singhal(1990), Routing in point-to-point delivery systems: formulations and solution heuristics, *Transportation Science, 24(4)*:245-260.

Cheung, W., Leung, L.C., and Y.M. Wong(2001), Strategic service network design for DHL Hong Kong, *Interfaces, 31(4)*:1-14.

Ko, H. J., Lee, H. J. and Ko, C. S.(2010). A study on the design of express courier service network based on the cut-off time adjustments. *International Journal of Innovative Computing, Information, and Control, 14:*159-177.

Lee, H.J. Ko, H. J., Sohn, Y. H. and Ko, C. S.(2010). Profit-based reconfiguration of express courier service network with multiple consolidation terminals. *International Journal of Logistics Systems and Management, 6(4):*442-457.

Goldberg, D.E.(1989). *Genetic algorithms in search, optimization and machine learning*, Reading, MA: Addison Wesley.

Gen, M. and Cheng, R. (2000). *Genetic Algorithms and Engineering Optimization*. Taylor John Wiley, New York, USA.

A SUPPLY CHAIN PERFORMANCE MEASUREMENT USING THE CONCEPT OF MOMENTUM

Min Kwan Cho[1] and Young Hae Lee[2*]
[1]LG CNS Entrue Consulting Partners
S.Korea
e-mail: mkwcho@lgcns.com

Department of Industrial and Management Engineering
Hanyang University
S.Korea
e-mail: yhlee@hanyang.ac.kr

Abstract: Most performance measures for the Supply Chain (SC) are concerned with either responsiveness or efficiency; therefore, only limited information is provided about the SC performance. Most SC performance is difficult to monitor in advance since these measures are based only on the results of SC operation. However, investigation of the SC state based on the response of supply activities to customer demands can enable SC performance analysis with respect to both responsiveness and efficiency. If customer demand, supply activity, and the SC are regarded as moving objects, it is described as a relative motion of objects that the response of supply activities to customer demands during the SC operation; furthermore, the concept of momentum can reasonably be used to analyze the response of supply activities to customer demands. Based on this physical analysis of the SC operation, our research presents a new SC performance measure called the Balancing Point (BP). The BP not only analyzes the SC performance with respect to both efficiency and responsiveness, but also considers forecasted customer demands and planned but not executed supply activities. In detail, it can estimate the present value of the future events over planning horizon related to the SC operation on SC performance. These features of BP enable the SC manager to detect unbalance between supply activities and customer demands. A numerical example is given to show the advantages of BP; also, suggestions are provided for further research on its use.

1. INTRODUCTION

The objective of Supply Chain Management (SCM) is to maximize profits of the overall SC, while satisfying customer requirement and minimizing the total SC operation cost. In order to achieve this objective, selection of appropriate performance measures is essential for SCM. The SC performance measure provides the basis for understanding the SC operation and information regarding the results of the SC operation; additionally, it influences behavior throughout the SC operation. These functions can be accomplished only when performance measures are based on the characteristics of SC.

SC characteristics are categorized into responsiveness and efficiency (Fisher et al., 1994 and Fisher, 1997). SC responsiveness is defined as the capability of the SC to completely satisfy customer demands; whereas SC efficiency refers to how economically the SC is operated. These characteristics, however, are contradictory to each other during SC operation. An example of this situation is the preferred inventory level of each characteristic. For responsiveness, in order to satisfy all customer demands without delay, a higher inventory level is preferred. Conversely, if a SC aimed at efficiency, it will regard inventory as a type of waste and try to decrease inventory level. Depending on the situation, a SC will range from focusing on being responsive to focusing on being efficient. As a result, the characteristic of a SC is determined by the trade-off between responsiveness and efficiency; therefore, it is necessary for the SC performance measure to deal with this trade-off.

Currently, the main performance measures applied in measuring SC performance are those for systems such as logistics, manufacturing, and inventory. In this case, the following problems are created (Stefan, 2000): a. Disconnect between strategy and performance, b. A biased focus on financial metrics, c. Too many isolated and incompatible measures in measuring SC performance, d. Lack of system thinking on the SC.

A main reason for these problems is that they do not concentrate on SC characteristics but on the specific function of the SC. This causes another critical problem; the trade-off between SC characteristics cannot be completely taken into account.

[*] corresponding author

For effective decision making on the SC operation, it is recommended that performance measures have the capability to evaluate SC performance in advance (Lapide, 1998). If the SC manager can recognize the SC state in advance, he can effectively coordinate SC operation. However, since performance measures almost always use the results of SC operation, such as inventory level or backlog, they only have after-the-fact information; thus, it is hard for the SC manager to make a proper decision about SC operation. Therefore, SC performance measure are required to estimate the value of a future event related to the SC operation, e.g., planned but not executed supply activities and forecasted customer demands on present SC performance(Melnyk *et al*, 2004)

The purpose of this research is to develop a new SC performance measure that not only handles both efficiency and responsiveness, but also estimates the SC state in advance.

2. SUPPLY AND DEMAND IN SC OPERATION

2.1 Three states of the SC with respect to response

All SC operations are coordinated in order to satisfy customers, which is practically achieved through the response to customer during SC operation. Here, the definition of response is: to provide customers with the products ordered by their due date. Thus, SC performance is directly affected by how the SC responds to customers. The response of supply activities to customer demands is important for an enterprise's outcome (Fisher *et al*., 1994).

In terms of response, various SC states can be classified into just three states: the demand dominating state, supply dominating state, and balanced state (Figure 1). The supply dominating state is the state in which the quantity provided by supply activities is greater than the quantity ordered by customers. In this state, surplus products remain until they are used in another response, which increases inventory cost in SC operation. Thus, these surplus products are regarded as a type of waste and must be decreased according to the efficiency viewpoint. When the ordered quantity is greater than the provided quantity, the SC will be in a demand dominating state. In this state, some customers are not satisfied; therefore, SC responsiveness will be decreased. If the provided quantity is ideally equal to the ordered quantity, the SC state will be in a balanced state. Conceptually, in this state, the response of the provided quantity completely offsets the ordered quantity, which results in all customers being satisfied and having no remaining products. Thus, the balanced state is the state in which responsiveness is harmonized with efficiency. Based on the information above, it is known that analyzing the SC state in terms of response can handle the trade-off between responsiveness and efficiency. Therefore, the response between supply activities and customer demands was used as the basis to develop the performance measure of this research

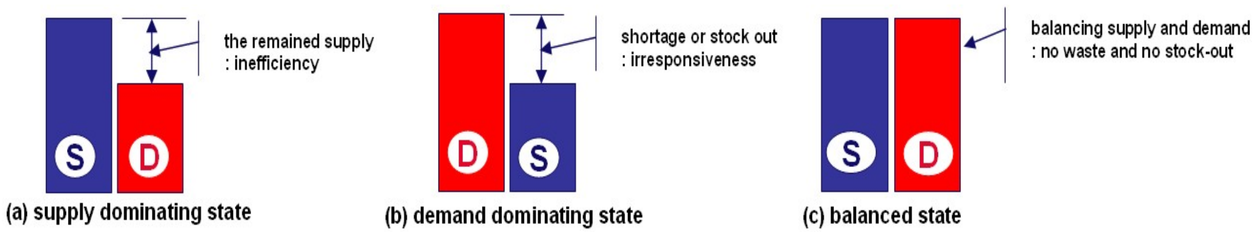


Figure 1. Three SC states in terms of the response of supply to demand

Developing the performance measure based on the response must be performed at the aggregate planning level. One reason is that supply activities in SC operation that respond to customer demand are coordinated in the aggregate planning level (Sunil and Peter, 2009). This is another reason that in SC aggregate planning level the uncertainty for the SC is not only relatively specified, but also effectively managed. Landeghem and Vanmaele (2002) classified uncertainty sources as the most appropriate planning level among strategic, tactical, and operational level to manage them based primarily on the time period over which they fluctuate. In their research, it is stated that there is not enough concrete information about the uncertainty at the strategic level and at operational level there is not sufficient time to react to the uncertainty. However, for aggregate planning at the tactical level, the information about the uncertainty is relatively concretized and sufficient time is available to decide upon appropriate countermeasures to prevent the SC operation from the disruptive impact of uncertainty.

3.2 Time and volume

Based on the meaning of responding in 3.1, the significance of responding during SC operation is examined in the domain of time and volume. The usage of a supply activity for responding can be analyzed with its execution date and the provided quantity by the supply activity. In detail, it is increased as the execution date gets closer to the timing of the response and the provided quantity increases. In terms of the time attribute, the actualization of a supply activity can be defined as how soon the supply activity will be executed by the SC in response to customer demands.

On the other hand, the influence of customer demand on response depends on the due date and the quantity ordered by the customer. It increases as its due date gets closer to the timing of the response and the ordered quantity gets larger. With respect to the time attribute, the urgency of customer demand refers to how soon supply activities must respond to the customer demand.

However, based on the rolling horizon mechanism, the present time of the SC operation will be progressed too, that is from the beginning to the end of planning horizon, and the SC will execute the supply activities and satisfy customer demands corresponding to the present time. Thus, the utilization of a supply activity and the influence of customer demand on responding will be changed as the present time of the SC operation advances. This follows what was previously mentioned, that the utilization of a supply activity and the influence of a customer demand on response can be formulated as functions with a time attribute, a volume attribute, and the present time of the SC operation as domains. In practice, Singhvi et al. (2004) developed an aggregate plan for the chemical industry using pinch analysis which handles time and quantity variables by plotting demand and production composites on a time versus material quantity plot.

3. DEVELOPING A NEW PERFORMANCE MEASURE FOR A SC

3.1 Physical analysis of SC operation and the concept of momentum

Based on the SC aggregate plan, a supply activity s supplies its provided quantity PQ_s at its execution date ED_s. Also, a customer demand d is fulfilled with its ordered quantity OQ_d at its due date DD_d during the planning horizon PH of SC operation. This means that the SC executes the supply activity, where execution date is the present time and satisfies the customer demand, where due date is the present time. Such SC operation can be analyzed as relative physical motion on the condition that a supply activity, customer demand, and SC are regarded respectively as moving objects (Figure 2).

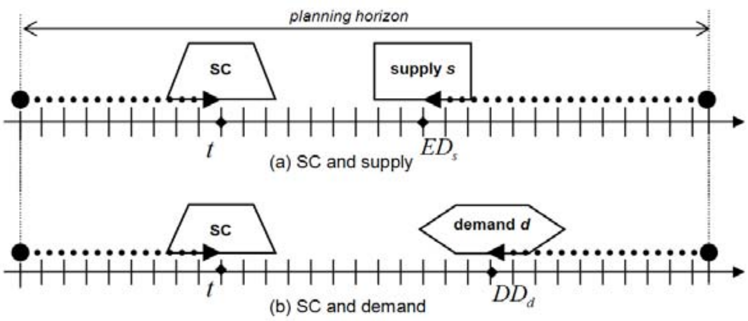


Figure 2. Customer demand, supply activity, and the SC during the planning horizon of SC operation

In detail, the SC could be described as the object moving from the beginning to the end of the planning horizon at time t in order to execute supply activities and satisfy customer demands. On the other hand, a supply activity s was explained as the moving object at ED_s from the end to the beginning of the planning horizon, so that the object may provide PQ_s to the SC only when the present time of the SC operation t reaches its execution time ED_s. Similarly, a customer demand d was described as the object which moved at DD_d from the end to the beginning of the planning horizon in order to receive OQ_d during SC operation. For this relative motion, ED_s, DD_d, and t were understood to be the destination of the motion of supply activity s, customer demand d, and SC, respectively. This physical analysis was based on the premise that the viewpoint domain of SC operation was converted from time to distance.

In physics, the concept of momentum is utilized when analyzing the interaction among objects in relative motion. The momentum **p** of a moving object is calculated using mass *m* and velocity **v** of the object as in the following equation:

$$\mathbf{p} = m \cdot \mathbf{v} = m \cdot \left(\Delta\mathbf{d} / \Delta\mathbf{t} \right) \tag{1}$$

In Eq. (1), Δt is the given time for the moving object and Δd is the changed position of the object during Δt. Thus, v is defined as the derivative of the position with respect to the given time. If the changed position is exchanged with the changed capability, the concept of velocity is useful to evaluate the changed level of capability during a given time. The concept of velocity, however, can be extended to the concept of relative velocity, which is the velocity of an object from an observers' viewpoint. The relative velocity, $v_{relative}$ of a moving object by an observer is measured using the velocity v_{object} of the moving object and the velocity $v_{relative}$ of the observer as follows:

$$v_{\text{relative}} = v_{\text{object}} - v_{\text{observer}} \tag{2}$$

The concept of relative velocity has an advantage in that it relatively measures the capability of an object in regards to the state or viewpoint of an observer. It was useful to represent the relative level of customer demand urgency and the relative level of actualization of a supply response activity with respect to the present time. In detail, if the SC at the present time was considered as the observer who watches a customer demand *d* at DD_d and a supply activity *s* at ED_s as moving objects, the urgency of the customer demand *d* and the actualization of a supply activity *s* could be relatively evaluated as considering the positions of the SC, the customer demand *d*, and the supply activity *s* changed on planning horizon in progress of time.

Initially, the relative time standard of the SC at *t* to *PH RTSC(t)* as perceived by an observer was evaluated using the concept of velocity as follows:

$$RTSC(t) = -\left(\frac{t}{PH} \right) \quad \left(0 \le t \le PH \right) \tag{3}$$

In Eq. (3), the concept of velocity was applied with the assumption that *PH* is Δt and the operated period of the SC from the beginning of *PH* to *t* is Δd. The reason that *RTSC (t)* has a negative value is that the direction the SC is moving is directly opposite to that of supply activities and customer demands. This means that it is the SC that the subject of responding to customer demands using supply activities during SC operation.

From Eq. (2), the relative level of urgency of a customer demand *d* to the SC at *t* over *PH* $RU_d(t)$ was evaluated with *RTSC(t)* and the level of urgency of demand *d* over *PH* U_d, the ratio of the position of a customer demand *d* to *PH*, in Eq. (4). Similarly, by Eq. (5), the relative level of actualization of a supply activity *s* to the SC at *t* over *PH* $RA_s(t)$ was calculated with *RTSC(t)* and the level of actualization of a supply activity *s* over *PH* A_s, the ratio of the position of supply activity *s* to *PH*. In Eq. (4) and (5), U_d and A_s were formulated with the concept of velocity under the assumption that the moving distance for the customer demand *d* and the supply activity *s* over the given time *PH* were $PH-DD_d$ and $PH-ED_s$, respectively in Figure 2.

$$\begin{aligned} RU_d(t) &= U_d - RTSC(t) \\ &= \begin{cases} 0 & , PH < DD_d \\ \dfrac{PH - DD_d}{PH} - \left(\dfrac{-t}{PH} \right) = \dfrac{PH - DD_d + t}{PH} & , t < DD_d \le PH \\ 1 & , DD_d < t \le PH \end{cases} \end{aligned} \tag{4}$$

$$\begin{aligned} RA_s(t) &= A_s - RTSC(t) \\ &= \begin{cases} 0 & , PH < ED_s \\ \dfrac{PH - ED_s}{PH} - \left(\dfrac{-t}{PH} \right) = \dfrac{PH - ED_s + t}{PH} & , t < ED_s \le PH \\ 1 & , ED_s < t \le PH \end{cases} \end{aligned} \tag{5}$$

The value of $RU_d(t)$ and $RA_s(t)$ were real numbers between 0 and 1, and calculated as dimensionless units since *PH,* DD_d, ED_s, and *t* were expressed in the same unit of time. The value 1 for $RU_d(t)$ means that the customer demand *d* is overdue or DD_d is t; such customer demands must be responded to immediately with available supply activities. On the other hand, the value 0 for $RU_d(t)$ represents that DD_d is beyond PH; therefore, it is not necessary for the SC to respond to the customer demand d. Thus, the nearer DD_d is to t, the nearer $RU_d(t)$ is to 1 and the farther DD_d is from t; then the nearer $RU_d(t)$ is to 0. The value 1 for $RA_s(t)$ represents that the supply activity s has already been executed or ED_s is t. Thus, the SC can immediately use the products by the supply activity s in response to customer demands. On the other hand, the value 0 for $RA_s(t)$ means that ED_s is out of $PH,$ so that the SC cannot use the supply activity s in response to demand. Thus it follows that the closer ED_s is to t, the closer $RA_s(t)$ is to 1 and the farther ED_s is from t; then the closer $RA_s(t)$ is to 0. From the above statements, it is known that $RU_d(t)$ and $RA_s(t)$ respectively determine the time value of a customer demand d and a supply activity s for response on the basis of the relationship between the time attribute and the present time.

In terms of the volume attribute, section 3.2 discusses how the utilization of a supply activity and the influence of a customer demand on response are respectively proportional to the provided quantity and the ordered quantity. However, in Eq. (1), the momentum of a moving object is proportional to its mass. Thus, the concept of momentum could represent the volume attribute related to the response with consideration of the provided quantity and the ordered quantity as the mass of each moving object. The above physical analysis on SC operation shows that the concept of momentum is appropriate in measuring SC performance in relation to the response. Therefore, the development of a new SC performance measure was done with the above physical analysis and the concept of momentum.

3.2 Using the concept of momentum to measure supply utilization in SC response

From Eq. (1) and (5), the supply utilization of a supply activity s for responding at t $SU_s(t)$ was derived as Eq. (6):

$$SU_s(t) = PQ_s \cdot RA_s(t) = \begin{cases} PQ_s \cdot \left(\dfrac{PH - ED_s + t}{PH} \right), & ED_s > t \\ PQ_s \cdot 1, & ED_s \le t \end{cases} \quad (6)$$

From Eq. (6), it is known that $SU_s(t)$ determines a type of weight on the supply activity s in responding with consideration of ED_s, PQ_s, and t over planning horizon. In Eq. (6), $SU_s(t)$ increases as t approaches ED_s and is proportioned to PQ_s. As the numerator and denominator of $RA_s(t)$ were commensurable, i.e. time, the measuring unit of $RA_s(t)$ was dimensionless. The measuring unit of PQ_s was the number of products and the units of $RA_s(t)$ was dimensionless, so that $SU_s(t)$ was measured as the number of products. In SC aggregate planning, SC manger can consider various supply activities, such as inventory, manufacturing during regular time, over time, and subcontracting for each period of an aggregate plan [3]. Thus, each utilization functions for them are as follows.

3.2.1 The supply utilization of inventory for responding

With respect to SC response, the inventory level of the SC at t $IL(t)$ is the product quantities that executed supply activities already delivered to the SC and that can be immediately provided to a customer at t. Thus, the relative level of actualization of $IL(t)$ $RAIL(t)$ is 1, intuitively. Without intuition, $RAIL(t)$ was derived using the level of actualization of $IL(t)$ $AIL(t)$ as follows:

$$RAIL(t) = AIL(t) - RTSC(t) = \frac{PH - t}{PH} - \left(\frac{-t}{PH} \right) = \frac{PH - t + t}{PH} = 1 \quad (7)$$

The measuring unit of $RAIL(t)$ was dimensionless since the numerator and denominator of $RA_s(t)$ were commensurable, that is, time. On the assumption that $IL(t)$ itself was the mass of an object, the supply utilization of $IL(t)$ for responding at t $UIL(t)$ was measured with Eq. (8). Since $IL(t)$ was given as a constant and $RAIL(t)$ was calculated as dimensionless, $UIL(t)$ was also measured as a constant and was the number of products.

$$UIL(t) = IL(t) \cdot RAIL(t) = IL(t) \cdot 1 \quad (8)$$

3.2.2 The supply utilization of manufacturing for responding within a regular time

Based on the SC aggregate plan, the workers of the SC will manufacture the assigned product quantities for a period p

EM_p. For manufacturing during regular time, the end time of a period p ET_p was considered as the execution date and EM_p was regarded as mass. Thus, the relative actualization of EM_p at t $RAEM_p(t)$, was derived from Eq.(5), and the supply utilization of EM_p for responding at t $UEM_p(t)$, was evaluated using Eq.(6) as follows:

$$UEM_p(t) = EM_p \cdot RAEM_p(t) = EM_p \cdot \left(\frac{PH - ET_p + t}{PH} \right) \tag{9}$$

However, EM_p was normally distributed with mean, μ_p[m], and variance, $(\sigma_p[\mathrm{m}])^2$ on the premise that the workers of the SC will maintain a stable manufacturing capability and try to meet EM_p within tolerance. As a result, $UEM_p(t)$ was given as a normal random variable in Eq. (10).

$$UEM_p(t) \sim N\left(RAEM_p(t) \cdot \mu_p[m], \left(RAEM_p(t) \right)^2 \cdot \left(\sigma_p[m] \right)^2 \right) \tag{10}$$

3.2.3 <u>The supply utilization of overtime for responding</u>

If necessary, the SC manager can consider the overtime for each period in SC aggregate planning. The workers of the SC will manufacture the assigned quantities of products during the overtime of a period p OT_p. Similarly, for such overtime, ET_p and OT_p were regarded as execution time and mass, respectively. Therefore, the relative actualization of OT_p at t $RAOT_p(t)$ was derived from Eq.(5), and the supply utilization of OT_p for responding at t $UOT_p(t)$ was calculated by Eq. (11).

$$UOT_p(t) = OT_p \cdot RAOT_p(t) = OT_p \cdot \left(\frac{PH - ET_p + t}{PH} \right) \tag{11}$$

However, manufacturing OT_p was executed by the manufacturing capability of the SC, so that OT_p was normally distributed with mean μ_p[o] and variance $(\sigma_p[\mathrm{o}])^2$. As a result, $UOT_p(t)$ was also derived as a normal random variable in Eq. (12).

$$UOT_p(t) \sim N\left(RAOT_p(t) \cdot \mu_p[o], \left(RAOT_p(t) \right)^2 \cdot \left(\sigma_p[o] \right)^2 \right) \tag{12}$$

3.2.4 <u>The supply utilization of subcontracting for responding</u>

Under the SC aggregate plan, the SC manager can subcontract a portion of manufacturing quantities to selected reliable suppliers. Selected suppliers deliver the subcontracted quantities for a period p SQ_p to the SC at the beginning of a time period p BT_p. For this subcontracting, the execution date was represented by BT_p and the mass was SQ_p. From Eq. (5) and (6), the supply utilization of SQ_p for responding at t $USQ_p(t)$ was derived using the relative actualization of SQ_p at t $RASQ_p(t)$, shown in Eq.(13). Assuming that selected suppliers deliver the subcontracted quantities exactly at their due date, SQ_p and $RASQ_p(t)$ were given as constant numbers. $USQ_p(t)$ was calculated as the number of products in Eq.(13), as SQ_p was given as the number of products and $RASQ_p(t)$ was dimensionless.

$$USQ_p(t) = SQ_p \cdot RASQ_p(t) = SQ_p \cdot \left(\frac{PH - BT_p + t}{PH} \right) \tag{13}$$

3.3 Measuring influence of demand on responding using the concept of momentum

Based on Eq. (1) and (4), the influence of a customer d for responding at t $DI_d(t)$ was formulated as follows:

$$\begin{aligned} DI_d(t) &= OQ_d \cdot RU_d(t) \\ &= \begin{cases} OQ_d \cdot \left(\frac{PH - DD_d + t}{PH} \right) & , DD_d > t \\ OQ_d \cdot 1 & , DD_d \le t \end{cases} \end{aligned} \tag{14}$$

Eq. (14) shows that $DI_d\ (t)$ determines the weight of a customer demand d for responding on the basis of OQ_d, DD_d, and t. In Eq. (14), $DI_d\ (t)$ is proportional to OQ_d and inversely proportional to the difference between t and DD_d. $RU_d\ (t)$ was dimensionless and OQ_d was given as the number of products, so that $DI_d\ (t)$ was evaluated as the number of products. In SC aggregate planning, the SC manager considers two types of demand: backlog and forecasted demand [3].

3.3.1 The influence of backlog demand on responding

The ordered quantity and due date of a backlog b, i.e. $OQBL_b$ and $DDBL_b$, are determined by customer. The influence of a backlog demand b on responding at t $IBL_b\ (t)$ was derived using the relative level of urgency of a backlog b on responding at t $RUBL_b\ (t)$ in Eq. (15).

From Eq. (4) and (14), the relative level of urgency of a backlog b on responding at t $RUBL_b\ (t)$ was derived, so that the demand influence of a backlog b on responding at t $IBL_b\ (t)$ was evaluated in Eq. (15). $OQBL_b$ was given as the number of products and $RUBL_b\ (t)$ was dimensionless, so that $IBL_b\ (t)$ was calculated as the number of products.

$$IBL_b(t) = OQBL_b \cdot RUBL_b(t) = \begin{cases} OQBL_b \cdot \left(\dfrac{PH - DDBL_b + t}{PH} \right), & DDBL_b > t \\ OQBL_b \cdot 1, & DDBL_b \le t \end{cases} \tag{15}$$

3.3.2 The influence of forecasted demand on responding

The SC manger forecasts the product quantity which will be ordered for each period in a planning horizon and plans supply activities through SC aggregate planning. For the forecasting by SC manager, the end time of a time period p ET_p and the forecasted demand of a period p FD_p were regarded as due date and mass, respectively. From Eq. (4) and (14), the relative level of urgency of FD_p $RUFD_p\ (t)$ was derived, so that the demand influence of FD_p for responding at t $IFD_p(t)$ was calculated as follows:

$$IFD_p(t) = FD_p \cdot RUFD_p(t) = FD_p \cdot \left(\frac{PH - ET_p + t}{PH} \right) \tag{16}$$

In Eq. (16), FD_p was generally given as a normal random variable with mean $\mu_p[d]$ and variance $(\sigma_p[d])^2$, so that $IFD_p(t)$ was a normal random variable as follows:

$$IFD_p(t) \sim N\left(RUFD_p(t) \cdot \mu_p[d], \left(RUFD_p(t)\right)^2 \cdot \left(\sigma_p[d]\right)^2 \right) \tag{17}$$

3.4 Balancing point

With respect to supply activities, customer demands affect the SC operation over planning horizon. Thus, it is available for measuring SC performance at t on the basis of the response of planned or executed supply activities to considered customer demands at t. Each supply activity in 4.2 has a common feature to provide products to SC, and the sum of each supply activity increases its size. This feature is similar to the sum of vectors in the same direction. Thus, on the assumption that each supply activity was independent of another one, the total supply utilization for responding at t $TSU(t)$ was estimated as the sum of utilization of all supply activities and represented as a normal distribution in Eq. (18). Here, np_t is the first period after t.

$$TSU(t) = UIL(t) + \sum_{p=np_t}^{all\ p} \left(UEM_p(t) + USQ_p(t) + UOT_p(t) \right) : TSU(t) \sim N(a,b) \tag{18}$$

$$a = UIL(t) + \sum_{p=np}^{P} \left(RAEM_p(t) \cdot \mu[m]_p + RAOT_p(t) \cdot \mu[o]_p + USC_p(t) \right), \quad b = \sum_{p=np_t}^{P} \left(\left(RUEM_p(t)\right)^2 \cdot \left(\sigma_p[m]\right)^2 + \left(RUOT_p(t)\right)^2 \cdot \left(\sigma_p[o]\right)^2 \right)$$

Another element for responding is the influence of customer demands. It is a common characteristic of customer demand to request products of the SC. Similarly, the sum of each demand influence enlarges its overall size, like the sum of vectors in the same direction. The total demand influence for responding at *t TDI(t)* was calculated as the sum of influence of all customer demands and represented as a normal distribution under the assumption that each customer demand is independent of one another as follows:

$$TDI(t) = \sum_{b=1}^{all b} IBL_b(t) + \sum_{p=np_t}^{all\ p} IFD_p(t) : TDI(t) \sim N(c,d) \tag{19}$$

$$c = \sum_{b=1}^{all\ b} IBL_b(t) + \sum_{p=np}^{all\ p} RUFD_p(t) \cdot \mu[d] \quad , \quad d = \sum_{p=np}^{all\ p} \left(RUFD_p(t)\right)^2 \cdot \left(\sigma_p[d]\right)^2$$

The response of supply activities to customer demands is similar to the sum of two vectors on opposite directions in a system. In detail, supply activities and customer demands are offset through responding, and then the difference in size between them is remained as a resultant vector. Thus, it is known that through responding the SC state is affected by the larger one in size between supply activities and customer demands over planning horizon as the direction of the system is determined by the one of the largest vector in the system through the sum of the two vectors.

Based on these facts, a new SC performance measure Balancing Point of the SC at t *BP(t)*, defined as the sum of *TDI(t)* and *TSU(t)*, is proposed. As mentioned above, supply activities have a positive direction and customer demands have a negative direction with respect to responding and they are independent of each other. Thus, *BP(t)* is measured as the number of products represented as a normal random variable in Eq. (20).

$$\begin{aligned} BP(t) &= TSU(t) + \left(-TDI(t)\right) = TSU(t) - TDI(t) \\ &= UIL(t) + \sum_{p=np_t}^{all\ p} \left(UEM_p(t) + USQ_p(t) + UOT_p(t)\right) - \sum_{b=1}^{all\ b} IBL_b(t) - \sum_{p=np_t}^{all\ p} IFD_p(t) : BP(t) \sim N(e,f) \end{aligned} \tag{20}$$

$$e = UIL(t) + \sum_{p=np}^{all\ p} \begin{pmatrix} RAEM_p(t) \cdot \mu[m]_p + RAOT_p(t) \cdot \mu[o]_p \\ + USQ_p(t) - RUFD_p(t) \cdot \mu[d] \end{pmatrix} - \sum_{b=1}^{all\ b} IBL_b(t) \quad ,$$

$$f = \sum_{p=np_t}^{all\ p} \left\{ \left(RAEM_p(t)\right)^2 \cdot \left(\sigma_p[m]\right)^2 + \left(RAOT_p(t)\right)^2 \cdot \left(\sigma_p[o]\right)^2 + \left(RUFD_p(t)\right)^2 \cdot \left(\sigma_p[d]\right)^2 \right\}$$

BP(t) can be applied in analyzing SC state as blow. When the SC is in a supply-dominating state, the mean of the *BP(t)* is located in a positive area. This means that the size of weighted supply activities are expected to be larger than the one of weighted customer demands over planning horizon with respect to the responding. Thus, surplus quantities may be frequently occurred in real SC operation. On the other hand, the mean of the *BP(t)* on the negative area represents that the size of weighted customer demands are expected to be larger than the one of weighted supply activities. It means that customer demands not satisfied on their due date may be often occurred in real SC operation. However, the mean of the *BP(t)* is ideally located at 0 only when the SC is in a balancing state. This ideal state is presented only when all customer demands are satisfied on their due date, as well as there are no remaining product quantities after responding over planning horizon.

The proposed measure, *BP(t)*, has advantages in measuring SC performance. First, *BP(t)* can handle both responsiveness and efficiency. Since *BP(t)* is based on the response between supply activities and customer demands, it can consider the trade-off between responsiveness and efficiency during SC operation. Thus, the SC manager can examine the characteristics of the SC ranging from responsiveness to efficiency as the value of *BP(t)*. Second, *BP(t)* has the ability to predict SC performance. In the measuring procedure of *BP(t)*, it is considered that the forecasted demand and planned but not executed supply activities as a normal random variable; therefore, *BP(t)* can estimate the performance of the SC with respect to responding in advance. Third, the *BP(t)* is a present time-based performance measure. Based on the concept of momentum, it can convert the value of the forecasted customer demands and planned supply activities for responding on the present SC performance. Thus, it can provide useful information for decision making on the present time of the SC operation. This feature is similar to net present value method in engineering economics, which converts all cash flows on the analysis period to an equivalent present value using a selected interest rate. It can evaluate the desirability of an alternative relative to the present or some base point in time.

4. NUMERICAL EXAMPLE

To show the advantages of *BP(t),* a numerical example is given. In the numerical example, it was assumed that the SC manager reliably forecasts customer demands and develops a SC aggregate plan with a planning horizon of 24 months using the Linear Programming (LP) model by Sunil and Peter (2009). This LP model is a generalized model to create a SC aggregate plan and determines the following SC operation parameters over the given time horizon: (1) Production quantity from regular time, over time, and subcontracted time, (2) Inventory held, (3) Backlog/Stock out quantity, (4)Workforce hired/laid off, (5) Manufacturing Capacity Increase/decrease. In order to represent real SC operation and get SC operation results, a simulation model was developed using ARENA. On running simulation model, products were produced based on the SC aggregate plan to respond the forecasted customer demands, and customer demands were occurred as forecasted in each period over planning horizon.

Based on forecasted customer demands, the SC aggregate plan, and the operation result from the simulation model (e.g. inventory level and backlog information), *UIL(t)*, $UEM_p(t)$, $UOT_p(t)$, $USQ_p\ (t)$, $IBL_b\ (t)$, and $IFD_p(t)$ were calculated at every month, *t* (*t*=1~24). By Eq. (18) and (19), *TSU(t)* and *TDI(t)* were evaluated, and then *BP(t)* was calculated using Eq. (20) in Table 1. The means of *TSU(t), TDI(t), and BP(t)* are shown in Figure 3

Table 1. The mean and variance of *BP(t)* during SC operation

t	*Mean*	*Variance*	*t*	*Mean*	*Variance*
0	492.042	4310.625	360	-1113.975	4557.283
30	473.923	4553.012	390	-952.756	4326.007
60	-102.786	4785.156	420	-1144.158	3995.130
90	-153.733	5003.220	450	-963.919	3620.712
120	-161.992	5203.264	480	-1133.717	3200.009
150	-461.433	5236.241	510	-918.750	2700.278
180	-622.947	5220.668	540	-1151.117	2236.224
210	-704.178	5220.938	570	-658.772	1777.656
240	-773.425	5177.257	600	-269.878	1296.719
270	-764.472	5144.835	630	-161.034	868.194
300	-989.498	4983.880	660	-179.927	543.333
330	-1046.911	4787.795	690	-123.517	252.561
			720	98.000	

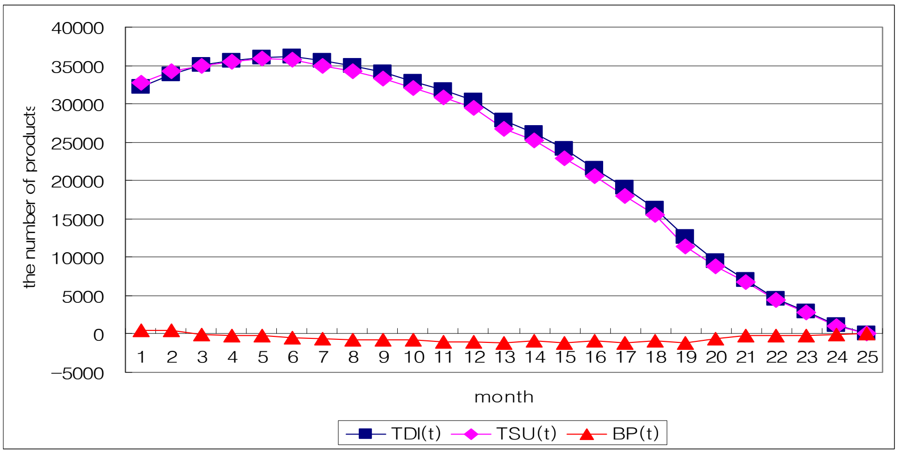


Figure 3. *TDI(t), TSU(t),* and *BP(t)* based on incorrect forecasting

In Figure 3, for the 24 months, the means of *TSU(t)* and *TDI(t)* are shown to be similar and the mean of *BP(t)* is close to 0 over planning horizon. More specifically, in rolling planning mechanism, whenever the SC manager evaluates *BP(t)* with respect to the remaining planning horizon at each time period, the SC is not only in a slightly demand-dominating state but also closely reaches a balanced state. The main reason for this situation is that customer demands over the planning horizon were reliably forecasted and the SC effectively responded to them with the developed SC aggregate plan. It means that all situations related to supply activities and customer demands were executed and occurred as planned and forecasted by the SC manager, and the SC operation has been optimized over planning horizon. However, as customer demands and the manufacturing capability of the SC were fluctuated within an allowable range, the acceptable difference between supply activities and customer demands occurs during the SC operation. These facts show that *BP(t)* is a reliable performance measure in analyzing SC state. In detail, for the above optimized SC states, *BP(t)* gives the information the state of SC has been in the balanced state over planning horizon.

5. CONCLUSION

In this research, a new SCM performance measure was developed Balancing Point, *BP(t)*, based on the concept of momentum. The *BP(t)* not only handles both responsiveness and efficiency in the trade-off relation by considering the response of supply activities to customer demands, but also predicts the SC performance by taking into account forecasted customer demands and planned but unexecuted supply activities as a normal random variable. In addition, it is a present time -based performance measure that converts the value of customer demands and supply activities to the present value with respect to responding. Due to these characteristics, *BP(t)* can enable the SC manager to actively make decisions about SC operation.

Basically *BP(t)* is developed on the level of SC aggregate planning so that it can be effectively applicable in S&OP(Sales & Operation) meeting, a type of consensus meeting. In this meeting, sales manager, production manager, and SC manager check whether products to be provided in order to satisfy demands without problem in overall viewpoint, and adjust supply plan or demand plan as necessary under the consensus among them. For above process, they need an indicator to show the expected overall SC status by applying supply plan and demand plan. In detail, if the indicator enables them to estimate whether shortage or oversupply, they must change supply plan or demand plan in order to solve such situation. As considering the advantages of *BP(t)* , it could be expected that *BP(t)* is enough to support such function. For example, if BP(t) is given in demand-dominating state, additional production must be considered in S&OP meeting. In oppostion case, the members of S&OP meeting must agree on the decreasing of production volume. However, if the *BP(t)* is wihtin accetable ranage with respect to balanced state, the present supply plan and demand plan must be maintained. As the mentioned above, it is known that BP(t) is an effctive tool in S&OP meeting.

Several future research areas could be considered using *BP(t)*. One of them would be the development of a replanning algorithm for the SC aggregate plan using *BP(t)*. It may be effective for the SC manager to decide to execute the replanning timing of the SC aggregate plan as the value of *BP(t)*. Another area of future research could be designing the framework to optimize SC operation using *BP(t)*. In a real-time operation, it is a very important to know when the system state has changed so that the operation can be adjusted as soon as possible. *BP(t)* is a time-value based measure as above mentioned, so that it can show the feedback result of a dispatch action on SC operation. This feature is useful to optimize SC operation in real time.

6. REFERENCES

Fisher, M. L. (1997). What is the right supply chain for your product? *Harvard Business Review*, *Mar.-Apr:* 105-116.

Fisher, M. L., Hammond, J. H., Obermeyer, W. R., and Ramen, A. (1994). Making Supply Meet Demand in an Uncertain World. *Harvard Business Review*, *May.-Jun.*: 83-93.

Landeghem, H. V. and Vanmaele. H. (2002). Robust planning: a new paradigm for demand chain planning. *Journal of Operations Management, 20:* 769–783.

Larry, L. (2000). What about measuring supply chain performance?, *Achieving Supply Chain Excellence Through Technology*, *2*: 287-297

Melnyk, S. A., Stewart, D. M., and Swink M. (2004). Metrics and performance measurement in operations management: dealing with the metrics maze. *Journal of Operations Management*, 22: 209–217.

Singhvi, A., Madhavan, K. P., and Shenoy, U. V. (2004). Pinch analysis for aggregate production planning in supply chains. *Computers and Chemical Engineering, 28:* 993-999.

Sunil, C. and Peter, M. (2009). *Supply Chain Management: Strategy, Planning, and Operation*, Prentice Hall, New Jersey,USA.

Stefan, H. (2000). A systems perspective on supply chain measurements. *International Journal of Physical Distribution & Logistics Management, 30, 10: 847-868.*

KEY PERFORMANCE INDICATORS FOR LOGISTICS AND GENERAL MANAGEMENT OF MANUFACTURERS IN PLASTIC INDUSTRY

Jirapan Liangrokapart

Department of Industrial Engineering
Mahidol University, Salaya Campus,
Nakornpathom, Thailand
Email: jliangr@yahoo.com

Abstract: The purpose of this study is to identify key performance indicators (KPIs) for manufacturers in plastic industry in Thailand using questionnaires survey. Four hundred and seventy eight questionnaires were sent to company managers in plastic industry by post. Thirty-nine questionnaires were completed and returned. The study found that customer perspective is the most important factor to measure performance, followed by internal perspective, financial perspective, and learning and development perspective, respectively. The logistics measures including 'on-time delivery percentage', 'cycle time', and 'order fulfillment cycle' are listed among the most important KPIs in plastic industry.

Keywords: Key Performance Indicators (KPIs), Plastic Industry, Balanced Scorecard, Logistics Measures

1. INTRODUCTION

This section introduces the importance of problems and describes the purpose of this study. In the past, as the number of manufacturing plants was limited and the industry competition was not so high, key performance indicators (KPIs) were not concerned as a tool for strategic management. In 1990s, when the competition was elevated, financial indicators were introduced in order to measure performance of the manufacturers. This has created a lot of critics, especially when the management had to make decisions based on financial indicators.

The financial indicators measure the manufacturers' income, expenditures, and surplus/deficit as well as related financial status. Thus, the indicators reflect only the past performance of the manufacturers. There is no way to know what they are doing well and whether or not they are going in the same direction as the manufacturers' strategic direction. For example, before economic crisis the financial performance of a company was so good. Based on their good financial performance, there was no signal that the company would be closed down soon due to the economic problems. Therefore, the company continued investing and expanding its businesses and infrastructures until it was bankrupted. If the company had considered other performance indicators before making strategic decisions, it would have known the problems, got some warning signals, and stopped investment and expansion. Then it could have been able to develop other strategic strategies to be survived.

In the recent growing of technology and the high competition, especially in Information and Communication Technology industry, organizations need to adjust themselves in order to make their competitive advantages. Most organizations adopt KPIs as a tool to measure and control organizations' efficiency as well as to identify their strengths and weaknesses. The balance of short-term and long-term measurement, internal and external measurement, as well as financial and non-financial measurement has been considered. Hence, the concept of balanced scorecard (BSC) has been initiated as a tool to evaluate organization performance with the four different perspectives: the financial, the internal business process, the customer, and the learning and growth.

Supply chain managers frequently speak of the need for "actionable" performance data which helps to form action plan and strategic decisions. What the supply chain managers want to know is what is being done well; what needs to be improved; where improvement is needed; and how to improve it (Liangrokapart and Leonard, 2002). These questions can be answered if the manufacturers have a well-designed set of performance indicators, which gives a clear and concise picture of the current situation of the manufacturers' performance and provides a way to understand the strengths and weaknesses. The continuous use of KPIs will help the organizations know their performance and find the ways for improvement with an ultimate aim for sustainable development.

The purpose of this study is to develop a list of appropriate key performance indicators (KPIs) for manufacturers in plastic industry in Thailand. A review of relevant literatures related to the concept of supply chain management, performance measurement of supply chains, and current situation in plastic industry is included in the next section. Then survey instrument, analysis of the survey results, and insights gained from the survey are discussed in detail followed by conclusions of the study and suggested avenues for further investigation.

2. LITERATURE REVIEW

In the study of key performance indicators for manufacturers in plastic industry, we had explored relevant publications to construct survey instruments. The literature review begins with the concept of Logistics and supply chain management, performance measurement of supply chain, and current situation in plastic industry.

2.1 Logistics and Supply Chain Management

Due to globalization and high competition, companies now focus on their own supply chains and try to use the advanced information and communication technology to continuously develop s supply chains starting from raw material sourcing, manufacturing of goods and services, delivering finished goods to inventory, and shipping to retailers or customers. A lot of studies have been conducted in the area of logistic and supply chain management.

Logistics is the management of the flow of goods, information and other resources between the point of origin and the point of consumption in order to meet the requirements of consumers (frequently, and originally, military organizations). Logistics involves the integration of information, transportation, inventory, warehousing, material handling, and packaging, and occasionally security. Logistics is a channel of the supply chain which adds the value of time and place utility (http://en.wikipedia.org/wiki/Logistics).

Simchi-Levi, et al. (2000) has defined supply chain management (SCM) as follows: "Supply chain management is a set of approaches utilized to efficiently integrate suppliers, manufacturers, warehouses, and stores, so that merchandise is produced and distributed at the right quantities, to the right locations, and at the right time, in order to minimize system-wide costs while satisfying service level requirements."

This definition leads to several observations. First, SCM takes into consideration every facility that has an impact on costs and plays a role in making the product conform to customer requirements; from supplier and manufacturing facilities through warehouses and distribution centers to retailers and stores. Second, the objective of SCM is to be efficient and cost-effective across the entire system. Finally, SCM encompasses firms' activities at many levels, from strategic level through tactical to operational level.

The use of supply chain management concept to increase market share for US competitive firms is challenging. Barriers to mange supply chain effectively include the failure to share information, fear of loss of control, lack of self awareness, enormity of supply chain, lack of supply chain satisfaction, lack of customer understanding, lack of understanding of supply chain, myopic strategies, and deficiency of mutuality (Benton and Maloni, 2005). Therefore, completed and accurate information is needed for developing good supply chain strategy. Organizations must understand their supply chain partners in all respects, including comprehension of the sources, imbalance, and consequences of power such that the most beneficial use (or disuse) of this power can be directed to achieve supply chain performance and member satisfaction.

2.2 Performance Measurement

Numerous authors have proposed performance measurement frameworks, which prescribe dimensions of performance to be monitored by organizations (Kaplan and Norton, 1996). Much of work focuses on the issue of designing measures and measurement systems. These two topics namely implementation of measurement systems and how to use them to manage business performance appear to be areas in which further research effort is required.

Folan and Browne (2005) supported that in the last 15 years, a lot of research has been done in the topic of performance measurement. Neely (1990) estimated that 3,615 articles on this topic were published during 1994-1996 and one book was launched every 2 weeks in 1996.

Neely (1999) identified the seven forces driving firms to focus on performance measurement which include the changing nature of work; increasing competition (globalization); specific improvement initiatives; national and international quality awards; changing organizational roles; changing external demands (growing power of customers); and the power of information technology. To measure performance of each partner in the supply chain effectively, a good performance measurement system should be implemented. Keebler, et al. (1999) proposed the characteristics for a good performance measurement system including is quantitative; is easy to understand ; encourages appropriate behavior; is visible; is defined and mutually understood; encompasses both outputs and inputs; measures only what is important; is multidimensional; uses economies of effort; and facilitates trust.

In addition, the concept of the balanced scorecard (BSC) has been introduced by Kaplan and Norton (1996) as a means to evaluate corporate performance from four different perspectives: the financial, the internal business process, the customer, and the learning and growth. The BSC framework for supply chain management includes translation of vision, communication and cooperation, business planning, and responsiveness and learning.

For effective supply chain management, the performance indicators should cover indicators for all management levels including strategic, tactical, and operational levels. The performance metrics proposed by Gunasekaran et al.

(2004) classified into each management level whereas Sharma and Bhagwat (2007) have classified supply chain metrics into the 4 perspectives of the BSC concept with an aim to evaluate supply chain performance comprehensively. They also recommended that the metrics should be reviewed and updated periodically. Another study of Sharma and Bhagwat (2007) has combined the concept of balanced scorecard (BSC) and analytical hierarchy process (AHP) in measuring performance of supply chains by ranking the importance and choosing appropriate metrics. The authors found that at the strategic level the customer perspective is the most important perspective, followed by internal business perspective. At tactical and operational levels, the innovation and learning perspective and the internal business perspective are the most important perspectives.

Several organizations could not maximize their supply chains' efficiency as they failed to develop appropriate performance measures and metrics which can be fully integrated with their supply chains. The further study of Gunasekaran et al. (2004) using questionnaires survey has shown a list of metrics for each production activity including plan, source, produce or assembly, and delivery. A seven-page questionnaire was divided into four basic sections: plan (including strategy), source/supply (order), produce (make/assemble), and delivery (to customer). The questionnaires were sent to a total of 150 large companies which were selected from a wide range of industry settings. Twenty-one questionnaires (14 percents) were completed and returned. Then they ranked the metrics from the most important metric to the least important metric for each production activity at each management level.

In addition to those studies mentioned earlier, Vaart and Donk (2008) have summarized that the use of supply chain metrics from previous studies might be difficult as different authors had used different supply chain measures at different levels. They have proposed the method to group the metrics into three categories including Attitude, Pattern, and Practice. Li, et al. (2005) has classified the metrics in the supply chain differently. They grouped the metrics into six categories including strategic supplier partnership, customer relationship, information sharing, information quality, internal lean practices, and postponement. All the six categories cover supply chain management activities from the beginning to the end including both the product and information flow in the supply chains.

2.3 Current Situation in Plastic Industry

Plastic industry is the downstream industry of petrochemical supply chain. In general, the manufacturers in Plastic industry use plastic granules, polymer or resin as raw materials to produce parts, products, synthetic fiber, or packaging. Products of Plastic industry include plastic packaging, house ware products, plastic toys and sport products, part or construction, electrical appliances and electronics, automobile parts, synthetic fibers, shoe parts, and other plastic products.

Lasschuit and Thijssen (2004) stated that in the early 1980s, progressive petroleum and chemical manufacturers have concerned their operational performance. The concept of performance measurement has become popular and the key performance areas (KPAs) and key performance indicators (KPIs) have been identified. Decisions and communications across the supply chain are still ineffective and delayed, because of offline erratic spreadsheets, misaligned performance drivers, functional barriers between departments and a lack of transparency across the supply chain. This leads to slow and inept day-to-day decisions that cost companies dearly in term of financial performance.

As manufacturers have moved toward a more integrated operation management function across the supply chain, it becomes necessary to measure the performance of the various parts of the supply chain on various dimensions, in a consistent way. However, many manufacturers are facing serious difficulties in implementing such supply chain-wide performance measurement systems that capture various performance dimensions at various levels in a consistent way. Lohman et al. (2004) have identified the causes of these difficulties as follows:

1. Decentralized, operational reporting history
2. Deficient insights in cohesion between metrics
3. Uncertainty of what to measure
4. Poor communication between reporters and users
5. Dispersed IT infrastructure

Among the causes of the difficulties, uncertainty of what to measure is one of the major concerns of supply chain managers. This study is aimed to identify the key performance indicators for the manufacturers in Plastic industry. The indicators will illustrate how their manufactures perform, what areas they are doing well, and what areas are needed for improvement.

3. METHODOLOGY AND ITS APPLICATIONS

The concept of key performance indicators (KPIs) is one of strategic management tools for effective and efficient management. Every organization is now concerned about its KPIs and the use of appropriate KPIs in measuring and comparing its performance with organizational standards and goals. The KPIs are used to evaluate the current performance of the organization and where the organization is heading to as well as suggest how to make improvement. The performance measurement should be in line with the organizations' goals and objectives with respect to each particular metric.

3.1 Identification and Selection of Appropriate KPIs

The study started by listing down all possible KPIs used by different previous studies. All the KPIs were classified into the four perspectives of the Balanced Scorecard. Then three experts in the Plastic industry were asked to shortlist KPIs which are frequently-used in the Plastic industry. Then questionnaires were developed using the experts-identified KPIs. The consensus list of KPIs is shown in Table 1 below.

Table 1. Expert-identified KPIs in the Plastic Industry

1. Financial Perspective	**2. Customer Perspective**
1. Cash Flow 2. Net profit 3. Revenue 4. Dividends 5. Debt 6. Accounts receivable turnover (Net income / average account receivable) 7. Return on investment (ROI) 8. Income growth level 9. Total assets 10. Market share 11. Economic value added (EVA)	1. Customer satisfaction 2. Customer complaints 3. Customer retention 4. Customer acquisition cost 5. Customer loyalty 6. Customer lost cost 7. Sales volume 8. Brand awareness 9. Number of customers 10. Reliability
3. Internal Process Perspective	**4. Learning and Growth Perspective**
1. Unit cost 2. Cycle time 3. Quality 4. Order fulfillment cycle 5. On-time delivery percentage 6. Defect Rate 7. Downtime 8. Warranty time 9. Frequency of returned purchases 10. Lead time 11. Inventory turnover 12. Time for new product introduction 13. Resource Utilization	1. Employee competency 2. Employee satisfaction 3. Motivation index 4. Training hours 5. Turnover rate 6. Quality of work environment 7. Lost time due to accidents 8. Reportable accidents 9. Employee suggestions 10. Information technology 11. Research and development expense

Most organizations expect to have high profitability, good liquidity management, fine risk management, and elevated interest coverage capability. These credentials make the organizations viable and successful. The measures of financial efficiency are important for the for-profit organizations. These measures tell the organizations whether or not their strategies work and how the strategic implementation affects the financial performance of the organizations.

Customer is an important aspect for organizations to do businesses. Measures with respect to customer perspective tell how organizations perform in creating value for the customers. This helps the organizations to find ways to improve their products and services in order to satisfy customers.

The internal process perspective is focused at how the organizations process internally to respond to customers and shareholders' need. Organizations should improve their internal process in the metrics that are directly related to the customers and shareholders concern. For example, customers want 'on-time delivery'; hence, the measure should be 'percentage of on-time delivery'. In case customers want 'quality'; the organizations have to think which internal process helps to make quality products or services. The measures should include 'response time', 'cost', and 'new product introduction'. Therefore, the organizations should understand what the customers' and shareholders' need and use proper measures to evaluate their performance. All logistic activities including Procurement, Technology Development, Inbound Logistics, Outbound Logistics and Operations are measured under the internal process perspective.

People, system, and technology are important factors for the learning and growth perspective. Measures in this respect should involve with people, system, and technology to identify competitive advantage for organizational success. Examples of performance indicators in this perspective include 'employee satisfaction', 'employee retention', and 'employee productivity'.

Each KPI has been defined clearly using the definition on the Wikipedia website <http://en.wikipedia.org/wiki/> and the adjustment by the industry experts as shown in Table 2.

3.2 Questionnaires

The chosen KPIs above were used in developing questionnaire survey to study the importance of each performance measures and metrics used in Plastic industry supply chain. For a common understanding of the respondents, each KPI is defined clearly in the questionnaires.

The questionnaires were divided into three basic sections including general information of the respondents; the importance of each KPI classified in the financial, customer, internal process, and learning and growth perspectives; and lastly an open-ended question asking the respondents to propose other KPIs for their organizations.

The questionnaires were sent to the plant managers by post with a covered letter stating the purpose of this study and asking their cooperation in responding to the questions and returning the questionnaires to the researcher. The target group of this study included the medium to large manufacturers which have more than 51 employees or a capital asset of more than 51 millions Baht as per the classification of manufacturers in manufacturing industry by the Ministry of Industry, Thailand. The study did not include small manufacturers, family businesses, which may lead to miss interpretation of the results due to the difference of the business sizes. Therefore, the total of 478 questionnaires was sent to the medium to large-sized plastic manufactures located in Thailand by post.

4. FINDINGS

4.1 Results from questionnaires survey

A total of thirty-nine questionnaires (or 8.15 percents response rate) was completed and returned. Section I of the questionnaires asked general information of the respondents including size of the plants, positions of the respondents, 100 percent Thai business or joint venture, type of business, and whether or not they have used KPIs in measuring their plants' performance.

Of the 39 respondents, the survey found that 57 percents were large-scale manufactures (more than 200 staff or a capital assets of more than 200 millions Baht) and the remaining 43 percents were medium-size manufacturers. Most of the respondents were plant managers; followed by plant engineers and production managers. A few respondents were human resources managers, accounting staff, deputy plant managers, administrative staff, and managing directors. Most of the manufactures are 100 percent belong to Thai businessmen and do not have any foreign partners. Most of the products (97.43 Percent) made for domestic sale only, not for export. Fifty-nine percents of the respondents operate as Business to Customer function, and the remaining 41 percents are Business to Business operations. Lastly, but not least, the survey found that 30 plants (76.92 percents) have identified their own KPIs and used them as an assessment tool to measure their performance.

In Section 2 of the questionnaires, the managers were asked to rate the importance of each KPI. A Likert scale from 1 to 5, in which 1 used to indicate very low importance and 5 is used to indicate very high important, were used to rate the importance of each KPI. Responses in the "5" range show that the dimension is very important whereas responses in the "4" range indicate that the dimension is somewhat important. Responses in the "3" range mean that the dimension is neutral important. Responses in the "2" range mean that the dimension is insignificant in the respondents' opinions, and responses in the "1" range judge that the dimension is very insignificant.

This section tells which KPIs are commonly used in the plant and how importance of each KPI. The average scores of each KPI are shown in Table 3. The survey showed that customer perspective (score of 4.33) is the most important perspective, followed by financial perspective (score of 4.25), internal process perspective (score of 4.10), and learning and growth perspective (score of 3.78) (see Table 4 and Figure 1).

Table 2. Definition of each KPI

1. **Financial Perspective**	**Definition**
1. Cash Flow	- the movement of cash into or out of a business, project, or financial product. It is usually measured during a specified, finite period of time.
2. Net profit	- the gross profit minus overheads minus interest payable plus/minus one off items for a given time period.
3. Revenue	- the income that a company receives from its normal business activities, usually from the sale of goods and services to customers.
4. Dividends	- the payment made by a corporation to its shareholder members. It is the portion of corporate profits paid out to stockholders.
5. Debt	- the amount which is owed. It is a means of using future purchasing power in the present before a summation has been earned.
6. Accounts receivable turnover	- a ratio measuring the number of times, on average, accounts receivables are collected during the period. It is calculated from (Net income / average account receivable)
7. Return on investment (ROI)	- a ratio of money gained or lost (whether realized or unrealized) on an investment relative to the amount of money invested.
8. Income growth level	- an increase in income over time.
9. Total assets	- Economic resources, tangible or intangible, which are capable of being owned or controlled to produce value and have economic value.
10. Market share	- the percentage or proportion of the total market volume. It can be expressed as a company's sales revenue (from that market) divided by the total sales revenue available in that market.
11. Economic value added (EVA)	- an estimate of economic profit, which can be determined, among other ways, by making corrective adjustments to GAAP accounting, including deducting the opportunity cost of equity capital.
2. **Customer Perspective**	**Definition**
1. Customer satisfaction	- a measure of how products and services supplied by a company meet or surpass customer expectation.
2. Customer complaints	- a customer’s expression of displeasure from low quality products/services or from other reasons.
3. Customer retention	- the activity that a selling organization undertakes in order to reduce customer defections. Successful customer retention starts with the first contact an organization has with a customer and continues throughout the entire lifetime of a relationship.
4. Customer acquisition cost	- the cost associated with convincing a consumer to buy your product or service, including research, marketing, and advertising costs.
5. Customer loyalty	- the behavior of repeat customers, as well as those that offer good ratings, reviews, or testimonials.
6. Customer lost cost	- the cost from customers leaving or not-buying the products/services.
7. Sales volume	- the total annual sales volume of a manufacturer.
8. Brand awareness	- a marketing concept that measures consumers' knowledge of a brand's existence. It refers to the proportion of consumers who know of the brand.
9. Number of customers	- the number of customers who buy products/services.
10. Reliability	- the ability of a manufacturer to perform and maintain its functions in routine circumstances, as well as hostile or unexpected circumstances.
3. **Internal Process Perspective**	**Definition**
1. Unit cost	- the cost per standard unit.
2. Cycle time	- the period required to complete one cycle of an operation; or to complete a function, job, or task from start to finish.
3. Quality	- the perception of the degree to which the product or service meets the customer's expectations.
4. Order fulfillment cycle	- the amount of time from customer authorization of a sales order to the customer receipt of product.
5. On-time delivery percentage	- the percentage calculated from dividing the orders delivered on-time by the total number of orders delivered.
6. Defect Rate	- the rate computed by: (number of defects) / (hours to do the review).
7. Downtime	- the period when a system is unavailable.
8. Warranty time	- the period for a collateral assurance or guarantee that the seller will provide for a specific remedy in the event of the products/services fail to meet the warranty.

Table 2. Definition of each KPI (continued)

3. Internal Process Perspective	**Definition**
9. Frequency of returned purchases	- the number of occurrences of returned purchases per unit time.
10. Lead time	- the period of time between the initiation of any process of production and the completion of that process.
11. Inventory turnover	- the number of times inventory is sold or used in a time period such as a year. It is equal to the cost of goods sold divided by the average inventory.
12. Time for new product Introduction	- the period of time for the complete process of bringing a new product or service to market.
13. Resource Utilization	- the percentage of time that resources are actually used, as compared with the total time that the resources are available for use.
4. Learning and Growth Perspective	**Definition**
1. Employee competency	- the employee's actual skills comparing to the list of skills and behaviors that are specific and well defined for a job.
2. Employee satisfaction	- the perception of the company employees, as well as their needs and expectations, that will be met through the company's quality system.
3. Motivation index	- Index of the motivation of the employees. Based on responses from a section in the employee survey.
4. Training hours	- the number of hours each employee gets training activities per year.
5. Turnover rate	- the rate at which an employer gains and loses employees.
6. Quality of work environment	- the extent to which employees can enhance their personal lives through their work environment and experiences.
7. Lost time due to accidents	- the period of time lost during accidents per year.
8. Reportable accidents	- the number of accidents reported in a year.
9. Employee suggestions	- the number of suggestions received from employees.
10. Information technology	- the number of computers provided to the employees.
11. Research and development expense	- the percentage of Research and development expenditure to the total operations expenditure.

Section 3 of the questionnaires is an open ended question asking the respondents to propose other important KPIs for their organizations. This section helps the author to know other important KPIs that the author had not included in the survey and the proposed KPIs will be used in future research. However, there was no other KPI proposed by the respondents. This can be implied that the list of KPIs cover all KPIs used in the plastic industry or the respondent hesitated to think about other KPIs.

4.2 Analysis of the results

The overall average score of all metric is 4.1. Twenty-two performance metrics have been determined as their scores are higher than overall average (see Table 5). The metrics are arranged from the higher scores to lower scores. It is obvious that among the financial indicators, the manufacturers consider 'revenue', 'net profit', and 'income growth level' as their performance evaluation tools, where as among the customer indicators, the manufacturers are interested in measuring 'customer satisfaction', 'customer retention', and 'reliability'. For internal process indicators, the respondents believe that 'defect rate', 'unit cost', and 'quality' indicators are among their most concern KPIs. The logistics measures including 'on-time delivery percentage', 'cycle time', and 'order fulfillment cycle' are also listed as the important KPIs in plastic industry. Last but not least, the 'quality of working environment' is ranked as the highest KPI in the learning and growth perspective.

Table 3. Average scores of each KPI

1. Financial Perspective	**Average score**	**2. Customer Perspective**	**Average score**
1. Cash Flow	4.42	1. Customer satisfaction	4.65
2. Net profit	4.54	2. Customer complaints	4.52
3. Revenue	4.71	3. Customer retention	4.63
4. Dividends	3.74	4. Customer acquisition cost	4.03
5. Debt	4.15	5. Customer loyalty	4.24
6. Accounts receivable turnover	4.35	6. Customer lost cost	4.37
(Net income / average		7. Sales volume	4.37
account receivable)		8. Brand awareness	3.93
7. Return on investment (ROI)	4.42	9. Number of customers	3.96
8. Income growth level	4.54	10. Reliability	4.62
9. Total assets	4.00		
10. Market share	4.12		
11. Economic value added (EVA)	3.80		
3. Internal Process Perspective	**Average score**	**4. Learning and Growth Perspective**	**Average score**
1. Unit cost	4.54	1. Employee competency	3.93
2. Cycle time	4.41	2. Employee satisfaction	3.86
3. Quality	4.52	3. Motivation index	3.88
4. Order fulfillment cycle	4.34	4. Training hours	3.77
5. On-time delivery percentage	4.52	5. Turnover rate	3.62
6. Defect Rate	4.65	6. Quality of work environment	4.10
7. Downtime	3.83	7. Lost time due to accidents	3.73
8. Warranty time	3.75	8. Reportable accidents	3.93
9. Frequency of returned	3.71	9. Employee suggestions	3.90
purchases		10. Information technology	3.59
10. Lead time	3.82	11. Research and development	3.28
11. Inventory turnover	3.63	expense	
12. Time for new product	3.61		
introduction	4.04		
13. Resource Utilization			

Table 4. Average scores of each BSC Perspective

BSC Perspective	**Average score**
1. Financial Perspective	4.25
2. Customer Perspective	4.33
3. Internal Process Perspective	4.1
4. Learning and Growth Perspective	3.78

Concerning all expert-identified metrics relating to Logistics management, the 'On time delivery percentage', 'Cycle time', 'Order fulfillment cycle' are selected as the first-tier metrics for Plastic manufacturers while the other metrics including 'Frequency of returned purchases', 'Lead time', 'Inventory turnover', 'Resource utilization', and 'Maintenance activities' are the second-tier metrics. It is obvious that the high-important Logistics metrics are measures somewhat relevant to customers and their expectation whereas the second-tier metrics measure manufacturers' logistics operations directly. For example, the 'On time delivery percentage' shows how good the manufacturer performed with respect to on-time delivery to the customers. The 'Cycle time' tells how long the manufacturer took to produce products/services to respond to the customers' need. Lastly, the 'Order fulfillment cycle' demonstrated the period from customer authorization of a sales order to the customer receipt of product.

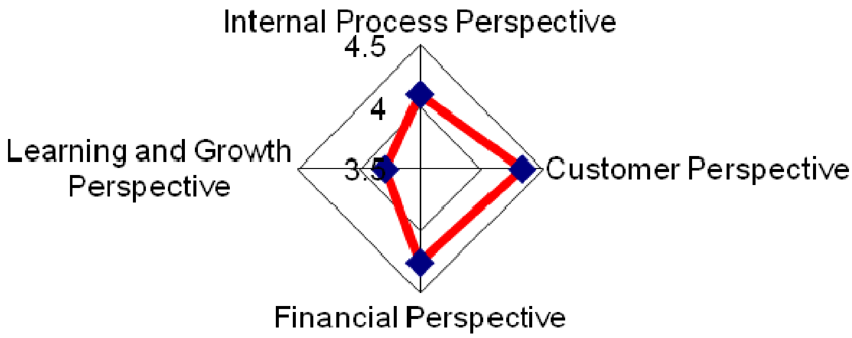


Figure 1. Importance of each perspective of BSC in Plastic Industry

Table 5. Important KPIs for each BSC perspective

1. Financial Perspective	**2. Customer Perspective**
1. Revenue 2. Net profit 3. Income growth level 4. Cash Flow 5. Return on investment (ROI) 6. Accounts receivable turnover (Net income / average account receivable) 7. Debt 8. Market share	1. Customer satisfaction 2. Customer retention 3. Reliability 4. Customer complaints 5. Customer lost cost 6. Sales volume 7. Customer loyalty
3. Internal Process Perspective	**4. Learning and Growth Perspective**
1. Defect Rate 2. Unit cost 3. Quality 4. On-time delivery percentage 5. Cycle time 6. Order fulfillment cycle	1. Quality of work environment

5. SUMMARY

Plastic industry is a for-profit industry where the ultimate goal of organizations in the industry is to maximize profit. As the industry is highly competitive, each manufacturer in the industry needs to know its performance and finds ways for improvement. Appropriate Key Performance Indicators have been identified in order to measure the manufacturers' performance and the Balanced Scorecard concept was applied to classify the Key Performance Indicators into 4 categories including Financial, Customer, Internal process, and Learning and growth perspectives. A total of 478 questionnaires were sent to all medium to large size manufacturers in plastic industry in Thailand and the return rate is

8.15 percents. The result of the survey shows a list of KPIs which are considered as important in the respondents' perception. From the four perspectives of BSC, the Customer perspective is the most important perspective, followed by Financial perspective, Internal process perspective, and Learning and growth perspective. Within each perspective, the important KPIs have been identified. It is hoped that the selected KPIs can be applied for manufacturers in plastic industry as a management tool to evaluate and monitor the manufacturers' performance quickly without a pile of needed information.

It is obvious that among the financial indicators, the manufacturers consider 'revenue', 'net profit', and 'income growth level' as their performance evaluation tools, where as among the customer indicators, the manufacturers are interested in measuring 'customer satisfaction', 'customer retention', and 'reliability'. For internal process indicators, the respondents believe that 'defect rate', 'unit cost', and 'quality' indicators are among their most concern KPIs. The logistics measures including 'on-time delivery percentage', 'cycle time', and 'order fulfillment cycle' are also listed as the important KPIs in plastic industry. Last but not least, the 'quality of working environment' is ranked as the highest KPI in the learning and growth perspective.

It is suggested that the methodology used in this study should be applied to identify important KPIs in other industries. After all, it should be applied to all industries in order to find common sets of important KPIs which will be benefit to the supply chain managers to measure performance of their organizations and know how good they are performing and which activities or functions need improvement. Moreover, besides the Balanced Scorecard, the industry-specific KPIs may be categorized in other different dimensions including by Financial or Non-financial aspect, by Internal or External factors, by short-term or Long-term indicators. Finally, the future research topic includes "Identifying how well the manufacturers perform comparing to the industry benchmarks and how to make improvement". Baseline data of each KPI from each manufacturer in Plastic industry should be collected. Then benchmarking the data with the industry leader for each KPI should be prepared. Finally, the research should suggest the avenues for improvement.

6. REFERENCES

Benton, W.C. and Maloni, M. (2005). The Influence of Power driven Buyer/Seller Relationships on Supply Chain Satisfaction. *Journal of Operations Management*, 23: 1-22.

Bhagwat, R. and Sharma, M.K. (2007). Performance Measurement of Supply Chain Management: A Balanced Scorecard Approach. *Computers & Industrial Engineering*, 53: 43-62.

Folan, P. and Browne, J. (2005). A Review of Performance Measurement: Towards Performance Management. *Computers in Industry*, 56: 663-680.

Gunasekaran, A. Pael, C., and Mcgaughey, R.E. (2004). A Framework for Supply Chain Performance Measurement. *International Journal of Production Economics*, 87: 333-347.

Gunasekaran, A, Williams, H.J., and McGoughey, R.E. (2005). Performance Measurement and Costing System in New Enterprise. *Technovation,* 25: 523-533.

Kaplan, R.S. and Norton, D.P. (1996). Using the Balanced Scorecard as a Strategic Management System. *Harvard Business Review*, 74: 75-85.

Keebler, J.S. (1999). *Keeping Score: Measuring the Business Value of Logistics in the Supply Chain.* Council of Logistics Management, Illinois.

Lasschuit, W. and Thijssen, N. (2004). Supporting Supply Chain Planning and Scheduling Decisions in the Oil and Chemical Industry. *Computers & Chemical Engineering*, 28: 863-870.

Li, S., Rao, S.S., Ragu-Nathan, T.S., and Ragu-Nathan, B. (2005). Development and Validation of a Measurement Instrument for Studying Supply Chain Management Practices. *Journal of Operations Management*, 23: 618-641.

Liangrokapart, J. and Leonard, M. (2002) Measuring and Enhancing the Performance of Closely-linked Decision Making Units in Supply Chains using Customer Satisfaction Data. *Proceedings of the IERC Conference*, Institute of Industrial Engineer, Orlando.

Lohman, C., Fortuin, L., and Wouters, M. (2004). Designing a Performance Measurement System: A Case Study. *European Journal of Operation Research,* 156: 267-286.

Neely, A. (1999). The Performance Measurement Revolution: Why now and what next?. *International Journal of Operations & Production Management,* 19: 205-228.

Sharma, M.K. and Bhagwat, R. (2007). An Integrated BSC-AHP Approach for Supply Chain Management Evaluation. *Measuring Business Excellence*, 11: 57-68.

Simchi-Levi, D., Kaminsky, P., and Simchi-Levi, E. (2000). *Designing and Managing the Supply Chain: Concept, Strategies, and Case Studies.* Mcgraw-Hill, Singapore.

Vaart, T.V.D. and Donk, D.P.V. (2008). A Critical Review of Survey-based Research in Supply Chain Integration. *International Journal of Production Economics*, 111: 42-55.

http://en.wikipedia.org/wiki/

BENEFITS ANALYSIS ON CIRCULAR CONSTRUCTION PRODUCTS SUPPLY CHAIN：CASE STUDY IN BEIJING

Zheng Kai[1], Ru Yihong[2*], Wang Ruijiang[2], and Zhu Yu[2]

[1] China Center for Industrial Security Research
Beijing Jiaotong University
No.3 Shangyuancun, xizhimenwai, Haidian District
Beijing, China
e-mail: zhengkai@bjtu.edu.cn

[2] School of Economics and Management
Beijing Jiaotong University
No.3 Shangyuancun, xizhimenwai, Haidian District
Beijing, China
e-mail: yhru@bjtu.edu.cn

Abstract: With the fast development of the construction industry in Beijing, the amount of construction solid waste moves upward, which causes serious environmental pollution. To solve these problems, It is urgent to establish the circular construction products supply chain (CCPSC). One of the key tasks unsolved is how to calculate the economic benefits and environmental benefits of CCPSC. Then in this paper, the property flow, material flow and cash flow models are set up for the CCPSC based. These models can be used to analysis the change of the economic benefits and the environmental benefits before and after the establishment of CCPSC.

1. INTRODUCTION

Being the important basic industry, the construction industry develops at a speed of 10% in recent years in Beijing. During the manufacturing process of construction products, on the one hand, much construction material (CM) has been consumed; on the other hand, the amount of construction solid waste (CSW) moves upward, which causes serious environmental pollution. To solve these increasingly serious problems, it is urgent to create a circular construction industry.

However, the current situation of construction products supply chain, which is a typical “linear logistics system” (Figure 1), can’t meet the demand of resources circulation. First, the majority of the CSW is buried in the disposing plants, and only a small part can be recycled, thus lots of land resources are occupied. Second, since the economic value of CSW is ignored, the construction companies can hardly benefit from recycling CSW, but they need to pay the treatment and logistics fee for CSW. Third, there are two types of logistics companies in the linear construction products supply chain (LCPSC), CM logistics companies and CSW logistics companies. They are supervised by different government sector, and the management objectives are inconsistent either. Therefore, such logistics facilities cannot be effectively integrated as distribution centers, vehicles and so on. To create a circular construction industry, we must create a new type of construction products supply chain that can substantially promote resources circulation, i.e. the circular construction product logistics system (Zheng *et al*., 2004).

CCPSC is a kind of resource-saving and environment-friendly supply chain which can integrate the construction material logistics and construction solid waste logistics, in order to maximize the resource circulation during the construction process (Figure 2). Some researches have discussed the scope, character and operational model of CCPSC, but little focus on how to calculate the economic benefits and environmental benefits of CCPSC(Zheng et al., 2007). Then in this paper, the property flow, material flow and cash flow models are set up for the current logistics system and the CCPSC, based on the MFA method.

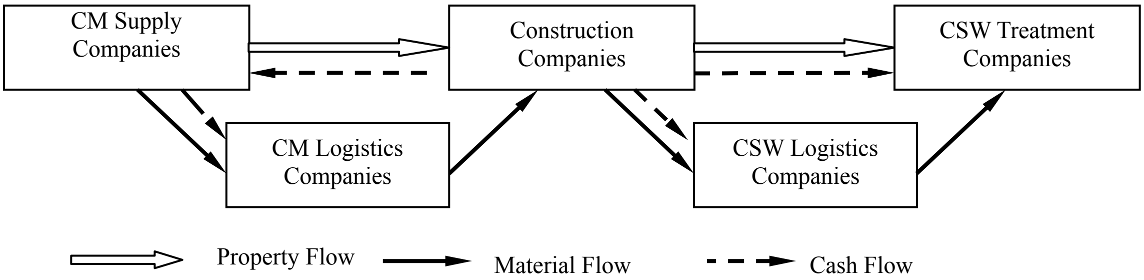


Figure 1. Linear Construction Products Supply Chain

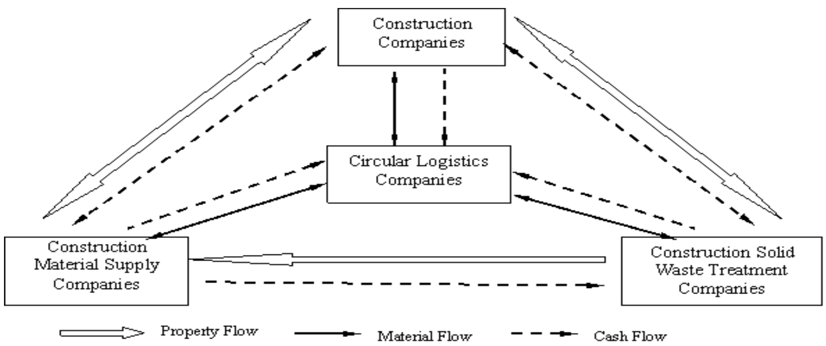


Figure 2. Circular Construction Products Supply Chain

2. METHODS

MFA method is a method of analyzing the flows of a material in a well-defined system. MFA recognizes that Material Throughput is required for all economic activities and asks whether the flow of materials is sustainable in terms of the environmental burden it creates. It accounts for all materials used in production and consumption, including the hidden flows, or ecological rucksack (Eurosta ,2001).

Besides material flow, there also exists cash flow in logistics system. In this paper, the material flow and cash flow will be mapped firstly based Figure 1, then the model for CCPSC will be set up in order to compare the economic benefits and environmental benefits. In theses models, some parameters can be used, as shown in Table 1, Table 2 and Table 3.

Table 1. Index and Sets

Indices and Sets	Illustration	Indices and Sets	Illustration
B	Superscript of CM supply companies	$l=1,2,\ldots,L$	Subscript of CM supply companies
E	Superscript of Construction companies	$j=1,2,\ldots,J$	Subscript of Construction companies
P	Superscript of CSW treatment companies	$k=1,2,\ldots,K$	Subscript of CSW treatment companies
W	Superscript of CM Logistics Companies	$m=1,2,\ldots,M$	Subscript of CM logistics companies
V	Superscript of CSW Logistics Companies	$n=1,2,\ldots,N$	Subscript of CSW logistics companies
U	Superscript of Circular Logistics Companies	$u=1,2,\ldots,U$	Subscript of Circular Logistics Companies
		$t=1,2,\ldots,T$	Periods

Table 2. Decision Variable List

Variable	Illustration	Variable	Illustration
x_{lt}^{b}	Mass of product purchased by company *l* in period *t*	u_{kt}^{p2}	Mass of recyclable materials sold by company *k* in period *t*
x_{ljt}^{be}	Mass of CW from company *l* to company *j* in period *t*	$x_{jj't}^{ee}$	Mass of CSW reused by company *l and* company *j'* in period *t*
u_{jt}^{e1}	Mass of reusable materials sold by company *j* in period *t*	x_{jlt}^{eb}	Mass of recyclable materials from company *j* to company *l* in period *t*
u_{jkt}^{ep}	Mass of CSW from company *j* to company k in period *t*	x_{klt}^{pb}	Mass of recyclable materials from company *k* to company *l* in period *t*
u_{kt}^{p}	Mass of CSW discharged by company *j* in period *t*	x_{kjt}^{pe}	Mass of recyclable materials from company *k* to company *j* in period *t*

Table 3. Parameter List

Parameter	Illustration	Parameter	Illustration
g_{jt}	Scale of construction products	c_{kt}^{p}	Unit price of recyclable materials sold by company *k* in period *t*
α	Productive efficiency of CM	p_1	Unit logistics fee of CM
β	Unit demand of CM	p_2	Unit logistics fee of CSW
γ	Unit product of CSW	p_3	Unit price of recyclable materials
c_{lt}^{b}	Unit price of product purchased by company *l* in period *t*	Q_{lt}^{b}	Maximum production capacity of company *l* in period *t*
c_{ljt}^{be}	Unit price of CM from company *l* to company *j* in period *t*	Q_{jt}^{e}	Maximum demand of company *j* in period *t*
c_{jkt}^{ep}	Unit price of CSW from company *j* to company *k* in period *t*	S_{jt}^{e}	Mass of CSW produced by company *j* in period *t*
c_{jt}^{e}	Unit price of reusable materials sold by company *j* in period *t*	Q_{kt}^{p}	Maximum disposal capacity of company *k* in period *t*

3. MODEL OF LCPSC

3.1 Flow Maps

Based on Figure 1, the material flow and cash flow of LCPSC are mapped, as shown in Figure 3 and Figure4.

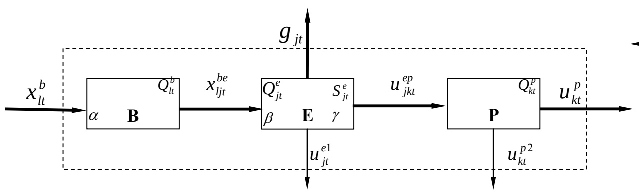


Figure 3. Material Flow of LCPSC

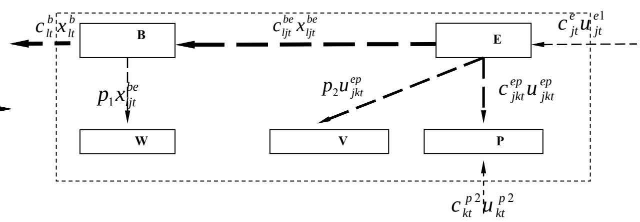


Figure 4. Cash Flow of LCPSC

3.2 Equations

In LCPSC, only the economic benefits are considered, so the objective for each company is to increase their sales or decrease their costs. We let F^b mean the benefit of CM supply companies; F^e the cost of Construction companies; F^p the benefit of CSW treatment companies; F^w the benefit of CM logistics companies; F^v the benefit of CSW logistics companies; the objective functions are as followed.

$$\max F^b = \sum_l \sum_j c_{ljt}^{be} x_{ljt}^{be} - \sum_l c_{lt}^{b} x_{lt}^{b} - \sum_l \sum_j p_1 x_{ljt}^{be} \tag{1}$$

$$\min F^e = \sum_l \sum_j c_{ljt}^{be} x_{ljt}^{be} + \sum_j \sum_k c_{jkt}^{ep} u_{jkt}^{ep} + \sum_j \sum_k p_2 u_{jkt}^{ep} \tag{2}$$

$$\max F^p = \sum_j \sum_k c_{jkt}^{ep} u_{jkt}^{ep} \tag{3}$$

$$\max F^m = \sum_l \sum_j p_1 x_{ljt}^{be} \tag{4}$$

$$\max F^n = \sum_j \sum_k p_2 u_{jkt}^{ep} \tag{5}$$

Constraints are as followed:

$$\alpha x_{lt}^{b} = \sum_j x_{ljt}^{be} \; . \tag{6}$$

$$\beta g_{jt} = Q_{jt}^{e} = \sum_l x_{ljt}^{be} \tag{7}$$

$$\gamma g_{jt} = S_{jt}^{e} = \sum_k u_{jkt}^{ep} \tag{8}$$

$$u_{kt}^{p} = \sum_j u_{jkt}^{ep} \tag{9}$$

$$x_{ljt}^{be} \le Q_{lt}^{b} \tag{10}$$

$$x_{ljt}^{be} \le Q_{jt}^{c} \tag{11}$$

4. MODEL OF CCPSC

4.1 Flow Maps

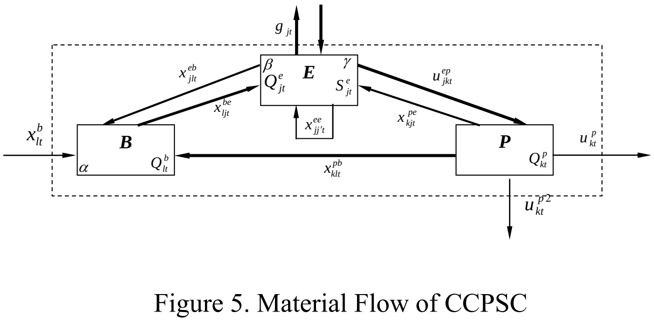


Figure 5. Material Flow of CCPSC

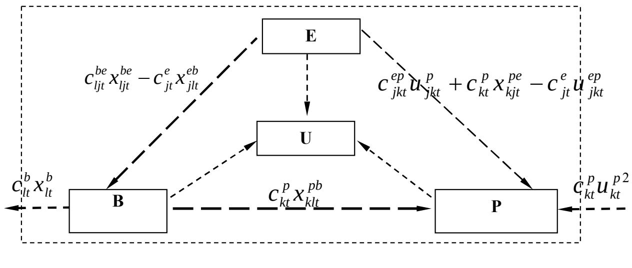


Figure 6. Cash Flow of CCPSC

Based on Figure 2, the material flow and cash flow of CCPSC are mapped, as shown in Figure 5 and Figure6.

4.2 Equations

In CCPSC, not only the economic benefits but also the environmental benefits are considered, so the objective for each company is to increase their sales or decrease their costs, on basis of increasing the recycling rate of CSW. We let R mean the recycling mass of the CSW, F^b benefit of CM supply companies; F^e the cost of Construction companies; F^p the benefit of CSW treatment companies; F^u the benefit of circular logistics companies; the objective functions are as followed.

$$\max R_1 = \sum_j \sum_l x_{jlt}^{eb} + \sum_k \sum_l x_{klt}^{pb} \tag{12}$$

$$\max F^b = \sum_l \sum_j c_{ljt}^{be} x_{ljt}^{be} - \sum_l \sum_j c_{jt}^{e} x_{jlt}^{eb} - \sum_l \sum_k c_{kt}^{p} x_{jkt}^{pb} - \sum_l c_{lt}^{b} x_{lt}^{b} - \sum_l \sum_j p_1 x_{ljt}^{be} \tag{13}$$

$$\min F^e = \sum_l \sum_j c_{ljt}^{be} x_{ljt}^{be} + \sum_j \sum_k c_{jkt}^{ep} u_{jkt}^{p} + \sum_j p_2 S_{jt}^{e} - \sum_j \sum_l c_{jt}^{e} x_{jlt}^{pb} - \sum_l \sum_j c_{jt}^{e} x_{jlt}^{eb} \tag{14}$$

$$\max F^p = \sum_l \sum_k c_{kt}^{p} x_{klt}^{pb} + \sum_j \sum_k c_{jkt}^{ep} u_{jkt}^{p} - \sum_j \sum_k c_{jt}^{e} u_{jkt}^{ep} - \sum_l \sum_k p_3 x_{klt}^{pb} \tag{15}$$

$$\max F^U = \sum_l \sum_j p_1 x_{ljt}^{be} + \sum_j \sum_k p_3 u_{jkt}^{ep} + \sum_l \sum_k p_3 x_{klt}^{pb} + \sum_l \sum_j p_2 x_{jlt}^{eb} \tag{16}$$

Constraints are as followed:

$$\alpha(x_{lt}^{b} + \sum_k x_{klt}^{pb} + \sum_j x_{jlt}^{eb}) = \sum_j x_{ljt}^{be} \tag{17}$$

$$\beta g_{jt} = Q_{jt}^{e} = \sum_l x_{ljt}^{be} \tag{18}$$

$$\gamma g_{jt} = S_{jt}^{e} = \sum_l x_{jlt}^{eb} + \sum_k u_{jkt}^{ep} \tag{19}$$

$$u_{kt}^{p} = \sum_j u_{jkt}^{ep} - \sum_l x_{klt}^{pb} \tag{20}$$

$$\sum_j x_{ljt}^{be} \le Q_{lt}^{b} \tag{21}$$

$$\sum_l x_{ljt}^{be} \le Q_{jt}^{e} \tag{22}$$

$$\sum_l x_{klt}^{pb} \le Q_{kt}^{p} \tag{23}$$

5. CASE STUDY

To test the validity of these models, 15 construction companies, 2 CM supply companies,1 CSW treatment companies l and 2 logistics companies are selected in Beijing as a case study. The number of parameters is confirmed through field investigations, as shown in Table 4. The calculate results are shown in Table 5.

Based on Table 5, the economic benefits and environmental benefits can be compared between LCPSC and CCPSC.

(1) The quantity of recyclable materials in CCPSC is about 27% more than that in LCPSC, which proves that CCPSC can create more environmental benefits.

(2) The gains or the costs of these companies are shown in Table 5, which proves that CCPSC can create more economic benefits.

Table 4. Parameter List for Case Study

Parameter	Number	Parameter	Number
g_{jt}	Random	c_{kt}^{p}	8.0
α	96%	p_1	1.5
β	1.0	p_2	1.5
γ	1.5	p_3	1.5
c_{lt}^{b}	10.0	Q_{lt}^{b}	10000
c_{ljt}^{be}	50.0	S_{jt}^{e}	1000
c_{jkt}^{ep}	1.5	Q_{kt}^{p}	40000000
c_{jt}^{e}	2.0		

Table 5. Solutions of the Models

Items		LCPSC	CCPSC
Environmental benefits	R	0.00	2.91
	R/S_{jt}^{e}	0.00	0.27
Economic benefits	CM supply companies(gain)	106.2	124.8
	Construction companies(cost)	171.8	164.2
	CSW treatment companies(gains)	16.1	17.9
	Logistics companies(gains)	20.3	21.5

6. CONCLUSION

The CCPSC is a kind of resource-saving and environment-friendly logistics system which can maximize the resource circulation during the construction phrase by means of integrating the construction material logistics and construction solid waste logistics. In this paper, the property flow, material flow and cash flow models are set up for the CCPSC based on the MFA method. It is proved that CCPSC can create more economic benefits and environmental benefits than LCPSC.

ACKNOWLEDGEMENTS

The authors are grateful to the generous support from China Postdoctoral Science Foundation and the China Fundamental Research Funds for the Central Universities.

7. REFERENCES

Eurosta (2001). *Economy- wide Material flow Accounts and Derived Indicator-A Methodological Guide*. Wiley, New York, USA.

Zheng Kai, Ru Yihong, Lin Zikui. (2004). Research on Reverse Construction Material Flow in Beijing. *Proceedings of 2004 International Conference on Logistics /MF,* Beijing Jiaotong University Press, Beijing, 2004, 564-571.

Zheng Kai, Ru Yihong, Yang Yongjie. (2007) Research on circular logistics system of the construction industry in Beijing. *Proceedings of International Conference on Construction and Real Estate Management*, Beijing, 2007

Session F2: IT

· **Day2: Sep. 16 (Thu.)**

· **Time: 10:40 - 12:00**

· **Chair: Bonghee Hong**

· **Room: Azalea & Lilac, 5F**

ENHANCED INTELLIGENCE OF WIRELESS SENSOR AND ACTUATOR NETWORKS IN LOGISTICS BY APPLICATION OF CONTEXT-AWARENESS

Vo Que Son[1*], Bernd-Ludwig Wenning[1], Carmelita Görg[1], and Andreas Timm-Giel[2]

[1] Communication Networks
University of Bremen
Otto-Hahn-Allee 1,
Bremen, Germany
e-mail: {son, wenn, cg}@comnets.uni-bremen.de

[2] Communication Networks
Hamburg University of Technology
Schwarzenbergstr. 95E
Hamburg, Germany
e-mail: timm-giel@tuhh.de

Abstract: Wireless Sensor and Actuator Networks (WSANs) are designed to operate with low power consumption and can react to changes of the physical environment by actuators. The optimization of efficient resource usage is of great importance in many research works. In real deployments, most applications of WSANs are configured to monitor and transmit the observed data periodically. This can lead to a waste of resources due to the duplication of traffic in the network if the physical environment does not change over time under normal conditions. In addition, most sensing applications only monitor the physical environment passively without making any reaction to unexpected changes of the environment. In this paper, a flexible context-aware model of WSANs is presented for use in logistic contexts to reduce unnecessary information overhead. Moreover, unexpected events happening in the environment can create warnings, e.g. by triggering a distributed local alarm system, activating an external controlling system or remotely alerting operators via emails or SMS. For flexibility of usage, this model is designed to be used at node level or central management system which is run at the sink. Logistics scenarios are simulated to illustrate the advantages of the proposed context-aware model.

1. INTRODUCTION

With the help of rapid development in digital electronics, sensor nodes with many advantages such as low-cost, low-power, distributed processing have been proposed for many applications such as environmental monitoring, asset tracking, and logistics. Sensor nodes are electronic devices which typically contain sensors, a microcontroller, a radio communication chip and other peripherals such as actuators and flash memory. They can both communicate with other nodes to form self-organizing WSANs and control the environmental surroundings by their actuators. Inheriting features from sensing, computation and communication technologies such as ad-hoc networking and distributed processing, WSANs allow telemetry, data collection and actuation, which can be suitable for logistic applications. This can contribute to a new logistic system which consists of intelligent logistic items that can autonomously control their transport processes (Son *et al.,* 10-2009).

A WSAN is context-aware if it can use context to provide relevant information to the user, to other sensors or also to itself (Hu aifeng *et al.*, 2005). A lot of research on context-awareness has focused on two main fields: routing and applications. The Privacy-Aware Location algorithm (Gruteser *et al.*, 2003) is proposed to prevent collection of privacy-sensitive data, other metrics (e.g. energy per packet, time to network partition) of Power-Aware routing are considered to prolong the lifetime of sensor nodes (Singh *et al.*, 1998). The remaining battery charge of nodes is also taken into account as a routing metric in this research. Environmental Monitoring Aware Routing (Wenning *et al.*, 2008) uses a multiplicative combination of environmental conditions and other context criteria for routing, which can be useful in disaster scenarios such as forest fires. A model of WSAN context-awareness (Marin-Perianu *et al.*, 2006) (Son *et al.*, 10-2009) is also proposed to use business rules in sensor nodes which can be applied for logistic transportation. However, all of the mentioned approaches are node centric and cannot be context-aware with respect to the whole deployed scenario.

Different from most of the above research works, which are studied in the area of routing, the model presented in this paper is designed in the application layer. The goals include some key points (Son *et al.*, 10-2009) as follows:

- Because of the dynamics in logistics, an adaptive proactive routing protocol could be used with this context-aware model to adapt to the changes of the network topology and support the mobility of logistics items. However, the context-aware application model should be independent from the routing layer.
- This model supports environmental and external conditions such as surrounding temperature, node positions, or connectivity between the sensor network and the infrastructure network (e.g. IP network).
- The context-awareness should be implemented in distributed nodes and at the central management system because the decisions to react to events from the environment have to be made in both the sensor nodes and the central manager.
- A rule-based context model should be used because this model is simple enough to be implemented in resource-limited sensor nodes and demands little computing and storage requirements (Marin-Perianu *et al.,* 2006).
- Supporting triggering or activating external sources such as alarm and HVAC (Heating-Ventilation, and Air-Conditioning) systems (HVAC).
- The contexts have to be easily programmable and the contexts are independent for each node.

In this paper, a context-aware model of WSANs is proposed to adapt to the collected environmental data. Moreover, this model does not eliminate the self-organization goal of WSANs by building the context model in the application layer while using the proactive routing protocol in (Son *et al.*, 08-2009). Besides the monitoring features based on contexts, this model also supports warning mechanisms such as triggering the alarm system.

This paper is structured in 5 sections: section I is the introduction to the scope of this paper, and related work. Context-awareness in logistics is discussed in section II. Section III describes the context-awareness model and underlying issues. Simulation results are shown in section IV. Finally, conclusions are given at the end of this paper.

2. CONTEXT-AWARENESS IN LOGISTICS

Although WSANs can generally be used in many fields, in the logistics domain WSANs can be applied in the following areas:

- *Transportation*: originating from the fact that WSNs have been deployed in smart containers (Jedermann *et al.,* 2007) to collect information about temperature, humidity or pressure of the good items, WSANs are becoming a promising technology for logistics. This requires sensor nodes to be attached to goods, shelves, Returnable Transport Items (RTIs) and containers to form a multi-hop sensor network. In addition, a container that is equipped with a WSAN must have a gateway to bridge the information from its WSAN to an external network (e.g. WLAN, UMTS) to make it accessible from outside (Son *et al.,* 10-2009).
- *Warehouse*: WSANs are also used on the warehouse floor where RTIs intended for a specific destination are grouped together. All rooms of the warehouse which store the logistic goods can be equipped with WSANs to allow users to monitor the physical environment of the warehouse completely. This requires a local alarm in each room and a HVAC system for alerting or controlling the environment changes.

Normally, all nodes in the network monitor and transmit data to a central management system via a gateway or a sink node.

In the real world, there are many contexts which need to be taken into account in deployments. However, in logistics, the context information can be obtained from some of the following sources (Son *et al*., 10-2009):

- *Physical environmental conditions* such as temperature, humidity, pressure. For example, deep-frozen food must be kept at a temperature of -18 °C or below. If the sensed temperature is higher than this threshold because of unknown reasons, the sensor nodes should send a message to users.
- *Security issues*: the goods packages or the container door can be secured by sensor nodes. If the security states are violated, the sensor nodes can trigger an alarm to warn about this event.
- *Position*: the location information is useful for the data collected, especially during the transportation. The absolute location can be used to determine the position of the container, or the relative location can provide the position of each item inside the container. This information can be useful for example when a problem is detected that requires an intervention at the specific location.
- *Long-distance connection*: a container using WSANs usually has a gateway to bridge the information between the WSANs inside the container and the outside network (e.g. WLAN, UMTS). In case there is no available connection to bridge the information, e.g. because of the signal coverage of the outside network, sensor nodes should temporarily store the data to prevent information loss.
- *Timing*: time is also a source of context. Sensors can be programmed to sample the environment at specific points of time.

3. MODEL OF CONTEXT-AWARENESS IN WSANS

In this section, a model is proposed to satisfy the previously mentioned issues. In our design the context-aware architecture is formed from the two following parts:

3.1 Context-awareness at sensor node level

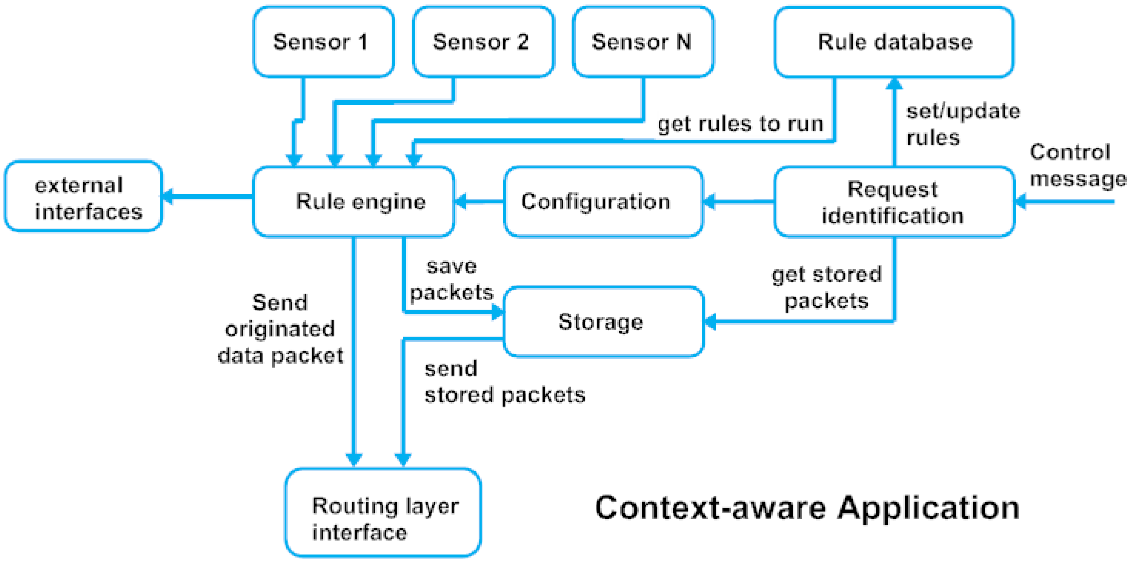


Figure 1. Model of context-aware application at node level

Figure 1 shows the architecture of the context-aware model at node level. It has main parts with the following functions:

- *Sensor*: all the internal and external sensors used by nodes for environment monitoring or other missions.
- *Rule database* contains a set of pre-defined rules which are run to trigger corresponding actions for the contexts.
- *Rule engine* runs all the rules retrieved from the Rule database to validate the data coming from sensors.
- *Configuration* contains the configuration for the operation of sensor nodes.
- *Request identification* is used to recognize the requests which can be used for updating rules, configuring the operation or retrieving the packets stored in Storage and other tasks.
- *Storage* is used to buffer the packets if necessary.
- *External interfaces* are used to communicate with external systems such as alarm or HVAC systems.
- *Routing layer interface* provides an interaction between the context-aware application and the routing layer to make it independent from the routing layer.

3.2 Context-awareness at central management level

When the context-aware model is applied in each node, it means that only that node can be aware of the environment surroundings. However, one or a few nodes cannot be aware of contexts of the entire network. For example, when the temperature of one room in a warehouse exceeds a predefined threshold, the sensor node which manages this room can know this event, but it cannot activate the central HVAC system of the entire warehouse since that node does not know the context conditions in other rooms. On the other hand, in case the contexts need a heavy computation, it is unwise to perform this computation at the node level.

Moreover, the context information sources do not come only from WSANs, for example the GPS signal identifying the global position of a sensor-equipped container is coming from outside the network.

Hence, if the context-awareness is also implemented at a central management system, the reactions to unexpected events are more accurate because information of all nodes in the network is available there. In order to support context-awareness at this level, the gateway (sink node) usually runs a management system which can interact with the WSAN for data collection and management. This system is more powerful when carrying out complicated algorithms. It also has extra interfaces such as controlling the HVAC system, warning users by sending an alert SMS (Short Message Service) or an e mail. These interfaces cannot be distributed to sensor nodes because of many reasons such as the expensiveness, the node size, and the power supply.

3.3 Context-aware rules at a node level

In order to support the context sources mentioned previously, a compact format for context rules is shown in Figure 2, which has been presented in a previous publication (Son *et al.,* 10-2009).

Rule ID (6)	Sensor Type (4)	Condition (4)	Min (16)	Max (16)	Action (4)	Next rule (6)

Figure 2. Format of context rules

In this context rule format, there are the following fields:

- *Rule ID* (6 bits): this unique number identifies the rule in the rule set. Hence, nodes can have many rules (up to 64 if they have enough memory) describing the actions for specific contexts.
- *Sensor Type* (4 bits): is used to determine which sensor will be used in this rule. The table 1 shows the values of this field.

Table 1. Values of Sensor Type field (Son *et al.,* 10-2009)

Sensor Type	Value	Meaning
NO_SENSOR	0	If the sensor node has no sensor
TEMPERATURE	1	If the testing sensor is Temperature
HUMIDITY	2	If the testing sensor is Humidity
LIGHT	3	If the testing sensor is Light
INVOLTAGE 4		If the testing sensor is Internal Voltage which measures the battery level
	5-15	Reserved for future use

- *Condition* (4 bits): the logical condition is used to check whether the rule applies. The pre-defined values of this field are shown in Table 2.

Table 2. Logical condition of rule (Son *et al.,* 10-2009)

Condition Valu	e	Meaning
BETWEEN	0	If the checking value is in range [Min..Max]
GREATER	1	If the checking data value of the checking packet is greater than Max
LESS	2	If the checking value of the checking packet is less than Min
OUT_OF	3	If the checking value of the packet is out of range [Min..Max]
GW_DISCONNECTED	4	If there is no connection between the Gateway and outside networks
GW_CONNECTED	5	If the Gateway is connected to any infrastructure network
TRIGGER_TIME	6	If the local time in a node is equal to the trigger time
	7-15	Reserved for future use

- *Min* (16 bits), *Max* (16 bits): the minimum and maximum values which are combined with the *Condition* field to form a logic condition. If the source of the context is the trigger time, both these two fields are used to describe a time stamp which indicates the time when the action in the rule is triggered.
- *Action* (4 bits): the corresponding action will be executed when the *Condition* is true. Pre-defined actions are shown in Table 3.

Table 3. Values of Action field (Son *et al.,* 10-2009)

Action Valu	e	Meaning
DO_NOTHING	0	Do not apply any actions with the current packet
SEND_PACKET	1	Send the checking packet to the next-hop
STORE_PACKET	2	Store current packet to memory
TRIGGER_ALARM	3	Trigger the local alarm when the context-aware rule is matched
ACTIVATE_HVAC 4		Activate the HVAC system (if available) when the context-aware rule is matched
	5-15	Reserved for future use

- *Next rule* (6 bits) is provided to construct logically linked chains to check the contexts. This helps to define a context from a set of rules by making a link chain between rules to shorten the processing time.

For example, a food package needs to be kept under the temperature lower than 4°C during the transportation. Any value of the monitored temperature greater than this threshold has to be reported. If sensor motes are used in this scenario, the context rule can be set as following:

IF TEMPERATURE GREATER THAN +4 °C THEN SEND_PACKET

Because sensor nodes usually only understand raw ADC values (12 bits or 16 bits), the human-understandable value (e.g. +4 °C) has to be converted to a raw ADC value by the context interpretation process.

The rule size is only 7 bytes; thus, sensor nodes with limited memory can still have many rules to describe contexts and related actions.

In comparison to the context-aware application model in (Son et al., 10-2009), this proposed architecture includes several improvements as follows:

- Instead of using the round robin technique to execute the rules, this proposal uses a logic link chain to create a rule sequence. This helps to define many contexts in a set of rules by linking them together.
- Both distributed and centralized context-awareness are used to enable the customization of context settings for the scenario and the creation of alert procedures.
- This architecture can be used in both WSNs and WSANs, where actuators are necessary to react to changes of the physical environment.

3.4 Context interpretation and programmability

The context interpretation should be performed at the central management system, which has the overall picture of the entire network. This also helps to reduce the computation at nodes that have limited resources. From the context descriptions, users have to define a set of rules which need to be executed in each sensor node. After that, the central management system will translate these rules into the rule format that the sensor node can understand using the numeric values in Table 1, 2, 3.

Because the contexts can change at any time, the rules must have the flexibility of being programmable to adapt to these changes. For the best flexibility, they can be programmed at any necessary time. Although the proposed model supports the rule programming at compilation time, it also supports remote programming of rules by sending commands (in control messages) to reconfigure the set of rules which describe the contexts. These control messages are implemented by using the dissemination technique (Levis et al., 2010), which is not only for programming context rules but also for other management purposes.

4. COMPUTER SIMULATION

4.1 Scenario 1: context-awareness of connection in transport logistics at node level

In this scenario, 20 packages equipped with sensor nodes are loaded into a container to form a WSAN. The container has a gateway to connect to the IP network and it is assumed to be transported to the destination in 3 hours. During the transportation, the connections between sensor nodes inside the container are always established thanks to the routing protocol (Son *et al.,* 08-2009). However, the connection between the gateway and the IP network may be disconnected in the second hour because of the coverage (shown in Figure 3). The other nodes in network are aware of the disconnection by a notification message sent from the gateway. In this context, the best solution is that nodes should transmit their sensed data when the connection between gateway and IP network is available. Otherwise, nodes store their packets in memory for the next transmission when the gateway connects to the IP network again. Nodes can be configured with normal or context operation mode. For evaluation, all data packets are monitored and logged at the gateway and at local nodes during 3 hours for analysis. The scenario is simulated in TOSSIM (Levis *et al.,* 2003) and TinyOS (TOS). All nodes except node 18 are configured to run in the normal operation mode, which report the sensed temperature in a data packet every 10 seconds. Only node 18 runs in context-aware mode with two following rules:

(1) if GW_DISCONNECTED then STORE_PACKET
(2) if GW_CONNECTED then SEND_PACKET

For comparison, the number of sent packets and received packets of node 13 which are captured at the gateway, is shown in Figure 3. It can be seen that node 13 lost all its packets (370 packets) during the time that the gateway disconnected because it is not configured to react with the surrounding context.

In contrast, Figure 4 shows the data collected from one node (node 18) and its local actions. One can see in the figure that the sensor node 18 only sends the packets when the connection is available in the first hour, afterwards it stores packets in local memory (if the memory is not full) in the second hour to avoid the loss of packets. This is indicated by the increase of received packets and the stored packets in the figure. When the gateway can reconnect to the IP network in the last hour, node 18 transmits all its sensed packets and its stored packets. Hence, the number of stored packets decreases in Figure 4 and the number of sent packets and received packets are the same which means that node 18 does not lose any packet during the simulation time.

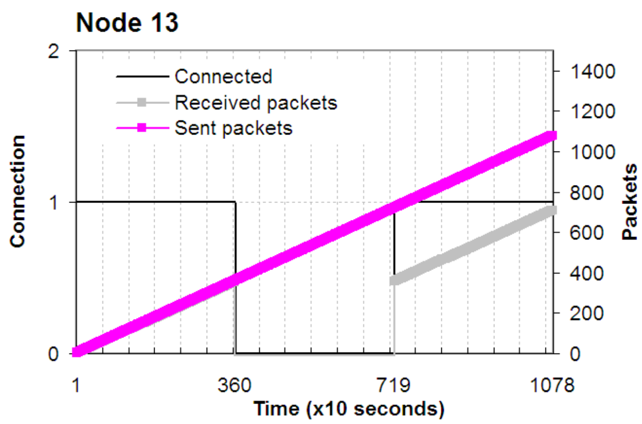


Figure 3. Connection and number of received, sent packets of node 13

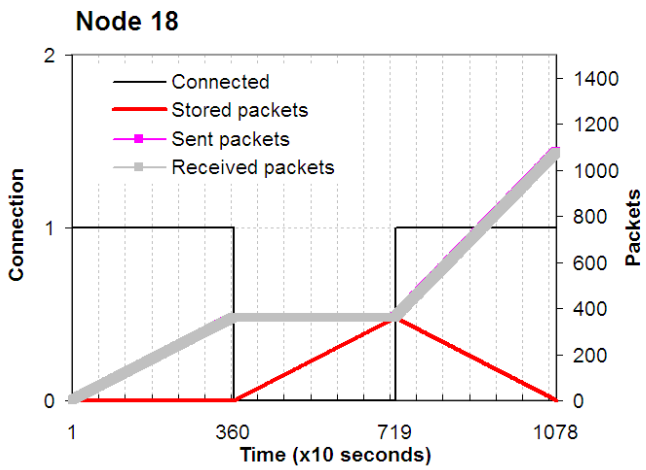


Figure 4. Connection and number of sent, receiving, and stored packets of node 18

4.2 Scenario 2: context-awareness of environment in a warehouse at node level

A food warehouse with the dimensions of 80m x 80m is used in this scenario for simulation. It has 16 equally-dimensioned square rooms as shown in Figure 5. Each room has a sensor node to monitor the environment inside and a local alarm system to warn in case of unexpected events such as high temperature. All sensor nodes (1..16) form a sensor network to report the sensed data to the sink node (node 0) using the routing protocol (Son *et al.,* 08-2009). Only node 0 is connected to a central HVAC system so that it can regulate the temperatures in all rooms of the warehouse. In the simulation, it is assumed that the room 8 contains frozen food which requires a temperature of -18 °C inside, while the others keep the temperature at +4 °C.

Because the details of controlling the HVAC system are out of this paper's scope, for simplicity, it is assumed that the HVAC sy stem is co ntrolled b y a d igital sig nal. If t his con trol signal is 1 , it m eans t hat th e HVAC system i s activated.

In the scenario, node 8 uses the following rule in its context-aware mode:

(3) if TEMPERATURE GREATER than -18 °C then TRIGGER_ALARM
(4) if TEMPERATURE GREATER than -18 °C then SEND_PACKET

Meanwhile, node 10 is configured as follows:

(5) if TEMPERATURE GREATER than +4 °C then TRIGGER_ALARM
(6) if TEMPERATURE GREATER than +4 °C then SEND_PACKET

The ab ove ru le (4) or (6) only p rocesses the o riginating pack ets of a lo cal n ode. It is no t app lied fo r the forward ing packets from the neighbors because the local node does not know the contexts of the neighbors.

The e nvironmental t emperatures i n rooms 8 a nd 10 ar e assum ed t o change fo llowing the template ch arts in Figure 6 and 7. These fi gures al so show that the local al arm signal is tri ggered imm ediately when the temperatures increase above the defined thresholds in both node 8 and node 10.

Besides the warm ing feature s, the context -awareness hel ps t o re duce the generated t raffic in each originating node and th e fo rwarding traffic in intermediate node s. Ta ble 4 sho ws t hat i n t he si mulated sce narios, the traffic generated by node 8 or 10 is lower than the traffic generated by those nodes which a re co nfigured in n ormal ope ration m ode. Theoretically, t he generated traffic reduction is equal to the rate of total time when the context rules are m atched over the t otal runn ing time of eac h node. T his fa ctor is also sh own i n t he Tabl e 4 wi th t he sam e val ues. He nce, t he de finition of context r ules i s ve ry i mportant bec ause i t c an a ffect t he e fficiency of t he network.

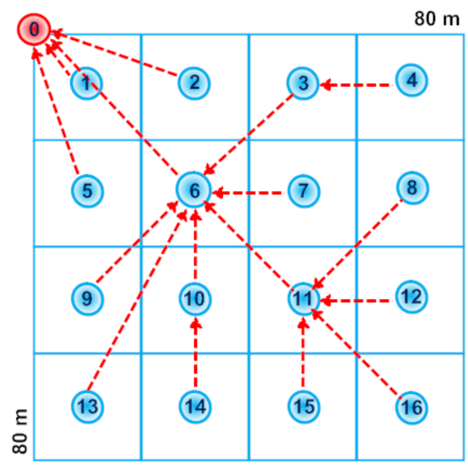


Figure 5. Layout of the warehouse and connectivity of nodes

Table 4. Traffic generated by nodes

Scenario Node		Number of transmitted packets (packets)		Reduction (%)
		Normal mode	Context mode	
2 1	0	1053	349	66.85
2 8		1056	351	66.76

4.3 Sce nario 3: c ontext-awareness at cen tral management system

Different from the two previous scenarios, in this scenario, the c ontext-awareness is applied at the ce ntral management syste m level. Actu ators a re us ed to react to the changes of environment from the central management system. Al l no des exce pt n ode 0 i n scenario 2 a re configured t o report t heir se nsed tem perature e very 10 seconds. T he pre set tem peratures in rooms are +3 °C. However, node 0 acti ng as th e cen tral manager will monitor all th e data from rooms (via sensor nodes) in this

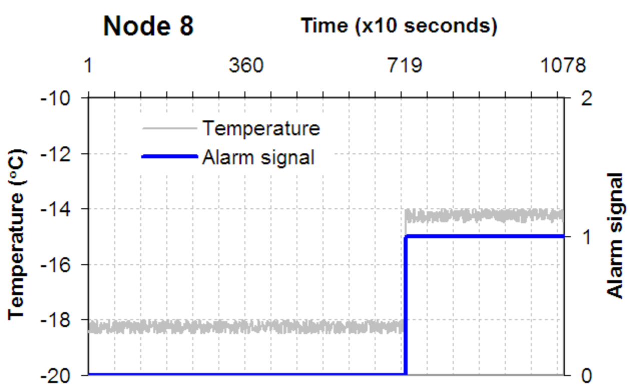


Figure 6. Change of temperature and alarm signal of node 8

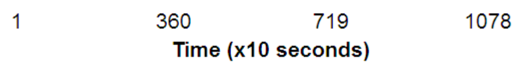


 Figure 7. Change of temperature and alarm signal of node 10

warehouse. If more than 10 rooms have temperatures higher than +6 °C, this node will trigger the HVAC system to adjust the temperatures in rooms of the warehouse.

Figure 8 shows the monitored temperatures of node 1, 2, 3, and 4 (other nodes have similar temperature curves). At the time of the 60th minute, the temperatures in the rooms suddenly increase due to unknown reasons, and exceed a value of +6°C some minutes later. Because node 0 can monitor all these changes, it triggers the HVAC signal to adjust the air-conditioning system to keep the temperatures in rooms at +3 °C again. When the temperatures go down, node 0 turns off the HVAC system since the monitored values are in the normal state again. This process is running in real time so the overall temperature is automatically kept at the temperature below the defined setting.

5. CONCLUSIONS AND OUTLOOK

With the proposed context-aware application model, it is believed that the logistic items are becoming more intelligent when they are equipped with WSANs. They not only know the conditions of surrounding environments, but also react to corresponding activities or communicate with other similar entities to send data. Combining the implementation of this model in distributed sensor nodes and at the central management system provides users the flexibility in describing contexts, collecting data and controlling the physical environment. Both logistics applications and others can utilize this model for many flexible targets to increase the communication as well as energy efficiency.

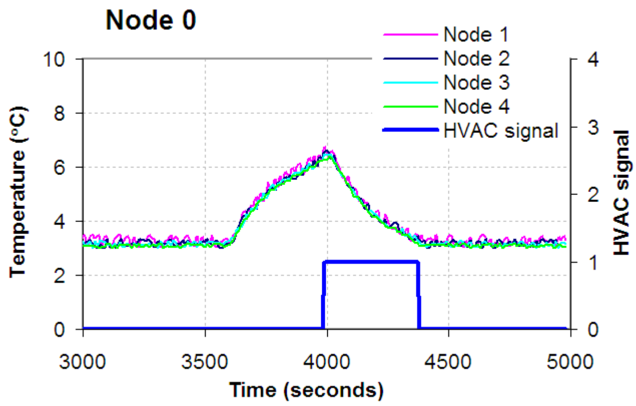


Figure 8. The change of temperature is reacted by HVAC control

Synergy between sensor nodes can be considered in the future to enhance the communication efficiency. Each sensor node not only uses its own context-awareness, but also interacts with others to improve the information quality. In addition, the context interpreter should support script files. Instead of configuring separated rules manually, the context interpreter can be improved so that it can translate a script file of context descriptions to a set of rules for each node.

ACKNOWLEDGEMENTS

This research is partially funded by the German Research Foundation (DFG) within the Collaborative Research Centre 637 "Autonomous Cooperating Logistic Processes: A Paradigm Shift and its Limitations" (SFB 637) at the University of Bremen, Germany.

6. REFERENCES

Huaifeng, Q., and Xingshe, Z. (2005). Context aware sensornet. *In Proceedings of the 3rd international workshop on Middleware for pervasive and ad-hoc computing*, pp. 1-7, Grenoble, France, November 2005.

Gruteser, M., Schelle, G., Jain, A., Han, R., and D. Grunwald (2003). Privacy-aware location sensor networks. *In Proceedings of the 9th conference on Hot Topics in Operating Systems*, pp. 28, Berkeley, USA, 2003.

Singh, S., Woo, M., and Raghavendra, C. S. (1998). Power-aware routing in mobile ad hoc networks. *In Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pp. 181-190, Texas, USA, 1998.

Wenning, B.L., Pesch, D., Timm-Giel, A., and Görg, C. (2008). Environmental Monitoring Aware Routing in Wireless Sensor Networks. *In Wireless and Mobile Networking. Proceedings of the IFIP joint conference on Mobile and Wireless Communications Networks (MWCN 2008) and Personal Wireless Communications (PWC 2008)*, pp. 5-16, Toulouse, France, September 2008.

Marin-Perianu, M., Meratnia, N., Lijding, M., and Havinga, P. (2006). Being Aware in Wireless Sensor Networks. *In Proceedings of the 15th IST Mobile & Wireless Communication Summit, Capturing Context and Context Aware Systems and Platforms Workshop*, Myconos, Greece, June 2006, Son, V.Q., Wenning, B.L., Timm-Giel, A., and Görg, C. (08-2009). A Model of Wireless Sensor Networks using Opportunistic Routing in Logistic Harbor Scenarios. *In Proceedings of 2nd International Conference on Dynamics in Logistics*, pp 214-223, Bremen, Germany, August 2009.

Jedermann, R., Behrens, C., Laur, R., and Lang, W. (2007). Intelligent Containers and Sensor Networks Approaches to apply Autonomous Cooperation on Systems with limited Resources. *In Understanding Autonomous Cooperation & Control in Logistics – The Impact on Management, Information and Communication and Material Flow*. Springer, pp. 365-392, Berlin, 2007.

Son, V.Q., Wenning, B.L., Timm-Giel, A., and Görg, C. (10-2009). A model of Wireless Sensor Networks using Context-awareness in Logistics Applications". *In Proceedings of the ITST2009 Conference*, pp. 2-7, Lille, France, October 2009.

Levis, P., Lee, N., Welsh, M., and Culler, D. (2003). TOSSIM: Accurate and Scalable Simulation of Entire TinyOS Applications. *In Proceedings of the First ACM Conference on Embedded Networked Sensor Systems*, pp. 126-137, Los Angeles, USA, November 2003.

Levis, P., Tolle, G (2010). Dissemination of Small Values. TEP118, http://www.tinyos.net/tinyos-2.x/doc

TOS: TinyOS . www.tinyos.net

HVAC: http://www.hvachome.net

INDEX CONSTRUCTION FOR RFID SENSOR TAG DATA IN RFID MIDDLEWARE

Yongxu Piao, Wooseok Ryu, and Bonghee Hong

Department of Computer Science and Engineering
Pusan National University
San 30, Jangjeon-dong, Geumjeong-gu
Busan, Republic of Korea
e-mail: {ywpark, wsryu, bhhong}@pusan.ac.kr

Abstract: With the development of RFID technique, it became a main applied technique for optimizing logistics management. However, some applications, like cold chain and chemical products management, issue new requirement for monitoring status of products. To address such requirements, RFID sensor tag which embedded a sensor device is developed by business companies. Such RFID sensor tag continuously measure and collect the products or environments' status information for logistics management. In this study, we focus on index construction for RFID sensor tag data in order to improve performance of query processing in RFID middleware. To efficiently construct index structure, the new data presentation for sensing data is proposed. It use a segment represent sensing information instead of a set of point. By this way, number of sensing is reduced. Another side, to improve the data representation accuracy, we use area threshold to limit number of data in one segment. Therefore, our sensing data representation makes a tradeoff between number and accuracy of sensing data.

1. INTRODUCTION

Logistics automation became the must-go road for efficient logistics management. To improve quality and efficiency of logistics, many newly adopted techniques have been devised to manage logistics asset. Among so many technologies, RFID (Radio Frequency IDentification) is used to automatically collect assets' information and locate the assets. With incessant growing of RFID and sensor, RFID sensor tag is developed by some business corporations. Except identification, it provides new functions to collect and measure assets status with embedded sensor device. For instance, some condition sensitive asset, chemical medicine and fresh vegetables, need to be kept in restrict conditions environments. So here RFID sensor tag attached on these assets in order to collect and monitor the status. Such sensing information helps user to determine the quality or freshness of product.

In logistics environment, a large number of RFID sensor tags are attached on the sensitive assets. During the asset's logistics lifecycle, the RFID sensor tag continuously measure and collect sensing information as the asset status records. Since, RFID sensor tag measures the status of asset in short time interval, and a lot of RFID sensor tags deployed in logistics system. This results in high volume sensing information being stored in RFID host repository. Many research efforts achieved on high volume data processing. GIS data management, Moving object database management etc al are related to high volume data management. No mater which research topics, the query processing over huge data repository is most important issue. It directly affects quality of system provided service to user. Luca. F and Ralf . H. G discussed the moving data model and structures for moving database. The discrete data model is suggested. This data representation can use be realistically used in a DBMS. Furthermore, several trajectory query processing technique have been presented. Three index structures, modified R-tree, STR-tree and TB-tree, are proposed for efficient trajectory query over moving object database. Many attentions are paid on query processing in moving object database. However, there no study on improvement of query processing performance for a large number of RFID sensor tag data.

In this study, we focus on improvement performance of query processing on RFID sensor tag data. In our target environment, a lot of RFID sensor tag deployed in logistics applications. RFID sensor tag data continuously append in RFID database. How to efficiently retrieve sensing information from RFID repository is important issue. The index structure is intuitionist way for improve query processing performance. To efficiently construct index for RFID sensor tag data, we suggest a new sensing data representation for RFID sensor tag data. In index structure, the sensing information is presented as a segment rather than a set of single points. This data representation can reduces number of indexing data. By this way, improve the query processing performance.

We organize our paper in five sections. Section 2 describes related work on index structure and data presentation on moving object database. Section 3 presents RFID sensor tag data properties and defines problem on indexing RFID sensor tag data. In section 4, we propose sensing data representation for efficient indexing RFID sensing data. At last, the summary of our work is given at section 5.

2. RELATED WORK

In this section, we introduce some related work on data index structure and moving object data representation. The basic index structure is explained. The data representation of moving object location is similar with sensing tag data. The segment data representation is an efficient way to reduce the number of the sensing data.

1.1 Data Index Structure

Index is auxiliary data structure that is used to improve the performance of data retrieve in database. Many index structures are developed by database society researchers. B-tree is useful index structure for one dimensional data. Several database softwares adopt this index structure to improve the performance of database. It uses a hierarchical index to minimize the number of disk reads. For multiple dimensional data, the R-tree is good choice. It is can be considered as multiple dimensional B-tree. Variants of R-tree are developed for different data model. Such as R+tree, R*tree and so on. For sensing data, the sensor device maybe has ability to measure multiple conditions at same time including temperature, humidity and illumination. Even only one condition need to be measured. The index structure still is multiple dimensional, since there are TagID and sensing time should be included in index structure.

1.2 Moving Object Data Representation

Moving object data representation is similar with sensing data. The moving object continuously generates location data. The update on location data is frequency. The data volume is also high. Sensing data have same properties. In moving object database, the segment data representation is used to present the movement of object. This segment is created on three dimensional X, Y and Time. X and Y is location information. Time presents temporal information for the specific location. The two endpoints of segment present the start location and stop location. During the time interval, consider the object move on straight line. If users want to know the object location in – between of endpoints of segment, take use of the endpoints and equation of the straight line to compute the approximate location of moving object. This representation greatly reduces the number of moving object location data. However, this data representation is not suitable for sensing data. Since, the moving object maybe always moves on straight line, the sensing value is not change according liner equation. The sensing value may change suddenly. This observation motivate us study on this topic and a efficient data representation for sensing data is proposed.

3. PROBLEM STATMENT

3.1 RFID Sensor Tag Data Properties

In this section, we describe properties of RFID sensor tag data. The problem on indexing RFID sensor tag data is presented. Main property of RFID sensor tag data is the high data volume, since a lot of RFID sensor tags are used in real applications and RFID sensor tag continuously collect sensing information in fixed short time interval in order to detect unexceptional situation in real time. We assume that RFID sensor tag measure ambient condition every 3 minutes and 5000 RFID sensor tags deployed in application environment. In this case, more than 2.4 millions sensing records are generated in one day. However, there may be more than 5000 RFID sensor tags deployed in real application. Unprecedented volume sensing information is generated.

Another property of RFID sensor tag data is bulk data appended. RFID sensor tag can not delver every condition measure to RFID host. The sensing information must be transformed to RFID host by RFID access point, i.e RFID reader. Only when RFID sensor tag arrive in range of RFID reader, the collected sensing information can be transformed to RFID host in form of sensing information set rather than single sensing information .

On view of RFID host, every RFID sensor tag data consists of Tag ID, ReaderID, ReadTime and a set of sensing information. Where, TagID is identification of RFID sensor tag, ReaderID is the accesspoint by which the RFID sensor tag data is transformed to RFID host. ReadTime specify when the RFID sensor tag data arrive RFID host. the set of sensing information is collection of measured condition records. Each record represents a once condition measure. It consists of sensing value and sensing time. For instance,

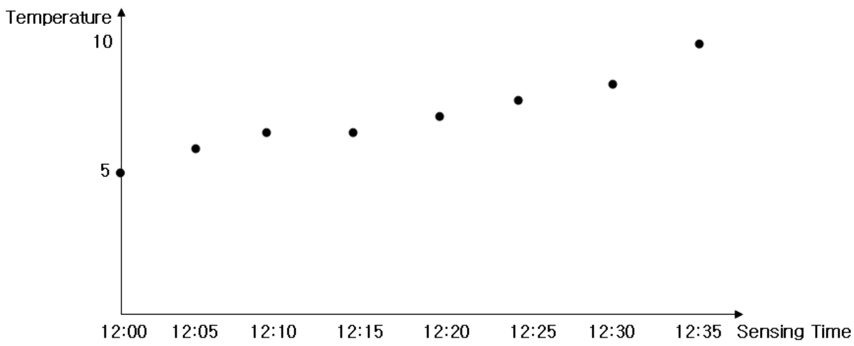


Figure 1. An Example of Sensing Data

Record <Temp erature, 5, 12:00:00> m eans th at t he specific RFID sen sor tag's tem perature co ndition is 5 at time 12:00:00. Figure 1 shows RFID sensor tag collected temperature information. In this example, This RFID sensor tag collected temperature information every 5 minutes from 12:00:00 to 12:40. Every point is once measure on temperature. This example only shows the sensing information on temperature dimension. Multiple dimension sensing information is also possible.

3.2 Problem on Indexing RFID Sensor Tag Data

Take use of RFID sen sor tag , th e asset con dition/statuses are co llected. Users retrieve RF ID host to g et in terested sensing information. RFID host should provide high quality service to end users. Performance of query processing over sensing information became critical factor . The intuitionist way for query processing i mprovement is index structure. Construct index structure using RFID sensor tag data. The naïve approach of index is that construct index structure on TagID, sensing value and sensing time three dimensions. Every sensing measure is point in three dimensions space. As shown in figure 2, point A present RFID sensor tag 1 ' temperature is 25 at time 13:00. However, there are so many sensing information need to insert into index structure. Every on RFID sensor tag data will bring frequency update on index structure. The update cost is very high, since RFID sensor tag data continuously arrive to RFID host. As result, index structure co nstructed by ra w se nsing i nformation c ause t wo problems. O ne i s t he high dat a v olume pr oblem. Another one is high update cost.

Such problems also happened in moving object database. To solve these problems, the segment representation is adopted in moving object database. Figure 2 (a) show an example of moving object data representation. The end points of segment are start point A and stop point B of moving object. That means the moving object move from location A' to location B' during T1 and T2. This representation greatly reduces the number of data. Accordingly, the update cost is low compare with raw location data. We try to use this approach on sensing data. In Figure (b), sensing information is represented as a segment AB. Bold line is real temperature values from T1 to T2. If users want to know the temperature value of Tc, th e resu lt co mputed fro m th is represen tation is v1. v2 is real te mperature value at T c. The re is great different between v1 and v2. T herefore, segment repre sentation is ap proximate re presentation. It is su itable for representation of moving object location. Since it only can exactly present start and end point. for RFID sensor tag data, accuracy of sensing information cannot be provided.

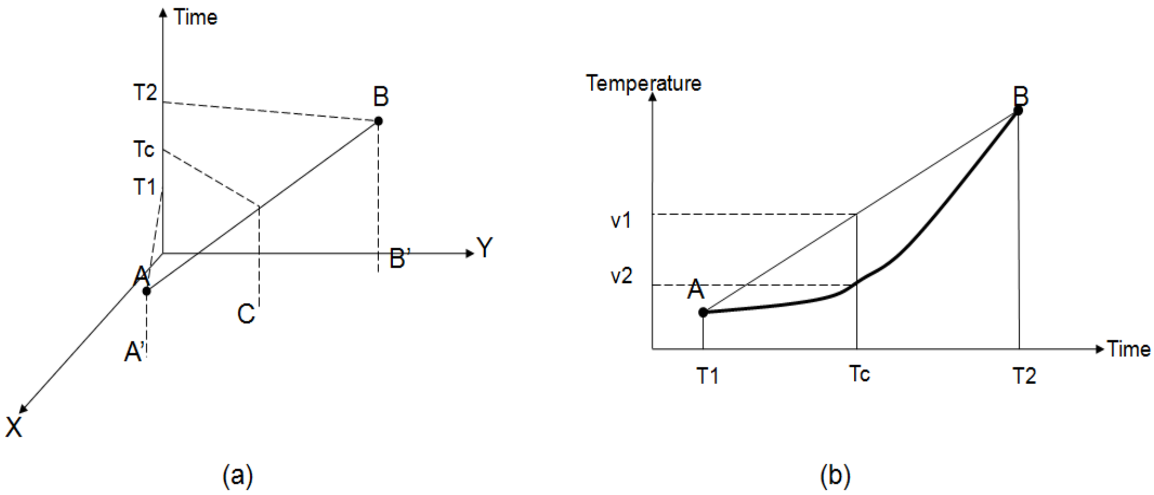


Figure 2. (a) Data Representation for Moving Object (b) Segment Representation for Sensing Data.

4. SENSING DATA REPRESENTATION

In this section, we explain our proposed sensing data representation. Basically, this representation is a segment representation instead of point data. It is different with segment representation of moving object data; proposed representation makes a tradeoff between number and accuracy of representation.

Let's discuss our sensing data representation in detail. As described in Section 2, the RFID sensor tag data have following properties; one is high data volume. Another property is that bulk sensing data is delivered to RFID host, since RFID sensor tag can connect to RFID host only when it is in range of RFID reader. In our approach, fist step is divide such once delivered bulk sensing information into sub data set and guarantee sensing data in every sub set is monotonic. Second step, limited area based division is executed. The reason of limited area is try to reduce the space of representation. Another import reason is aim to reduce different between lowest sensing value and high sensing value. Figure 3 is a sensing data representation example. Figure 3 (a) show the temperature data set collected by RFID sensor tag. The first step of our approach is presented in Figure 3 (b). The raw sensing data is divided into three parts. Each part is monotonic. The monotonicity makes sure that the endpoints of segment are the lowest and highest value in the segment. The sensing value and time can be considered as an equation of a straight line. Next step is dividing sensing data into small segment. We pre-define a representation area limitation for sensing data collection. This

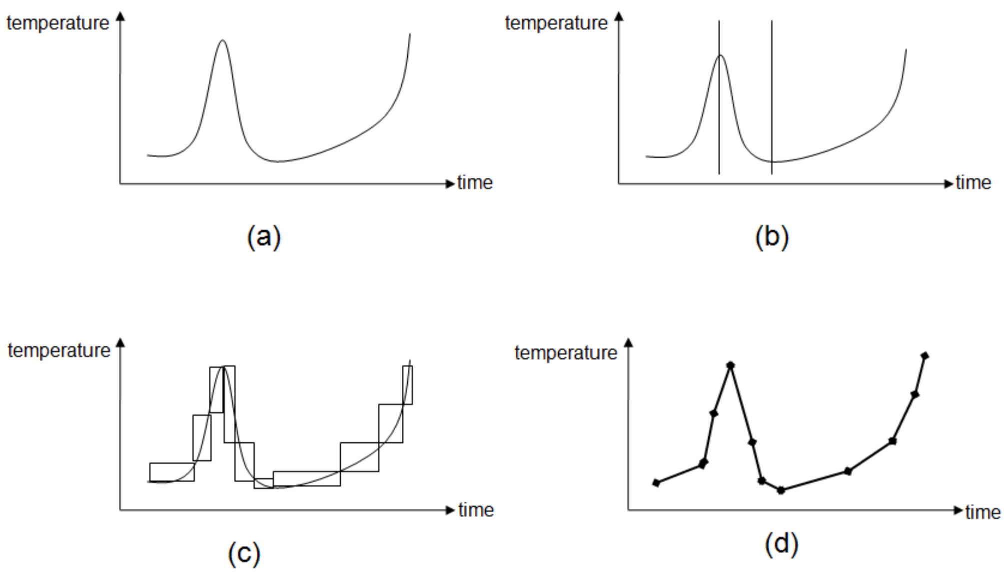


Figure 3. (a) Raw Sensing Data Collected by RFID Sensor Tag, (b) Divide Sensing Data into Monotonic Part (c) Limited Area based Partition (d) Final Data Representation for Sensing Data

representation area is sensing time* sensing value. Every sensing data representation represents a set of sensing data. The data set is from start point until sensing data representation area arrives to pre-define area. For instance, a segment represents a set of sensing data. The minimum sensing value is 5 and maximum sensing value is 10. The sensing data collect from 12 :00 to 13:00. The segment representation area is (10 -5)*(13:00-12:00) = 300 . If 300 less than the pre-defined area limitation, then add next sensing information into this segment until the representation area arrive to area limitation. The final data representation of sensing data is show in Figure 3 (d). All segments are used to construct index structure.

Our proposed data representation not only reduces the number of sensing data but also provides the accuracy of sensing information.

5. CONCLUSION

In this study, we aim to construct a high performance index structure for RFID sensing tag data. The problems on sensing data indexing are high volume of data and frequency update. That cause low performance of index structure and high update cost. To solve these problems, we suggest a new sensing data representation. This representation uses segment to present a set of sensing information rather than a set of point. Segment representation reduces the number of sensing data. However, it brings approximate data representation. To improve the accuracy of data representation, we limit the number of data in one segment using area limitation. It can provide approximate accuracy for RFID sensor tag data representation.

The proposed data representation makes a tradeoff between number of data and accuracy of representation. However, this data representation only is suitable for gradually changed data. For jitter sensing data, the data change suddenly, it is difficult to find monotonicity of sensing data. Our proposed representation is not suitable for such sensing data.

ACKNOWLEDGEMENTS

This work was supported by the Grant of Korean Ministry of Education, Science and Technology" (The Regional Core Research Program/Institute of Logistics Information Technology)

6. REFERENCES

Antonm. G, (1984). R-Trees: A DYNAMIC INDEX STUCTURE FOR SPATIAL SEARCHING. *Proc. ACM SIGMOD International Conference of Management of Data, 47-57.*

Dieter. P, Christian S. J and Yannis. T. (2000). Novel Approaches in Query Processing for Moving Object Trajectories. *Proceedings of the 26th International Conference on Very Large Data Base, 395-406.*

Douglas. C. (1979). The Ubiquitous B-Tree. *ACM Comput. Surv. 11(2): 121-137*

Luca. F. and Ralf. H. G. (2000). A Data Model and Data Structures for Moving Objects Database. *Proceedings of the 2000 ACM SIGMOD international conference on Management of data,* 319-330.

Xiao. F. M and Z himing (2003). D. DSTTMOD: A Discrete Spatio-Temporal Trajectory Based Moving Object Database System. http://idke.ruc.edu.cn/

AGENT-BASED CONTROL ARCHITECTURE TO IMPROVE THE PERFORMANCE OF RFID

Hae Kyung Lee[1], Taioun Kim[2*]

[1]Department of Industrial & Management Engineering
Kyungsung University
314-79 Kaeyeon-3 dong, Nam-gu
Busan, Korea
e-mail: hshklee72@ks.ac.kr [1]
twkim@ks.ac.kr [2*]

Abstract: To remain competitive in business and to be quick responsive in the warehouse and supply chain, the use of RFID (radio frequency identification) has been increasing in many industry areas. RFID can identify multiple objects simultaneously as well as identifying individual objects respectively. Some limitations of RFID still remain in the low recognition rate and the sensitive response according to the material type and its environment. Much efforts have been tried to enhance the recognition rate and to be more robust. Researches on tag design change, antenna angle, search angle and signal intensity etc correspond to the efforts.

The purpose of this paper is to propose an agent-based control system utilizing database, knowledge base and ontology in order to improve the performance of RFID. First, ontology is constructed for the environmental factors to be used as a knowledge base. Second, agent-based RFID manager middleware is created to increase the recognition rate in changing situations. Finally, a thorough experiment is performed to verify the method and contributions using the prototype.

1. INTRODUCTION

The radio frequency identification (RFID) technology allows remote identification of objects using radio signal, thus without the need for line-of-sight or manual positioning of each item. With the rapid development of RFID technology and its application, we expect a brighter future in the object identification and control. The major advantage of RFID technology over the barcode is that the RFID system allows detection of multiple items simultaneously as they pass through a reader field. Additionally, each physical object has its unique ID (even two products of the same type have two different IDs) enabling to precisely track and monitor the position of each individually labeled product piece.

An ontology is a formal representation of the knowledge by a set of concepts within a domain and the relationships between those concepts. It is used to reason about the properties of that domain, and may be used to describe the domain. An ontology provides a shared vocabulary, which can be used to model a domain — that is, the type of objects and/or concepts that exist, and their properties and relations. The focus of the ontology lies on the representation of the RFID domain. The ontology can act as a model for exploring various aspects of the domain. Since part of the ontology deals with the classification of RFID applications and requirements, it can also be used for supporting decisions on the suitability of particular RFID tags for different applications.

Previous researches of RFID are included regarding RFID device, middleware, agent, ontology and industrial applications. Jung (2009) presented a detail survey result for the RFID smart packaging guideline. In this government-funded research, they showed recognition result of RFID according to material type, packaging type, tag type, reader type and tag location etc. This research can provide a good guideline for applying RFID in various industry applications. Cho et al. (2007) proposed a RFID system modeling on manufacturing operation strategy. This research proposes steps to apply RFID system according to the manufacturing strategy in order to realize ubiquitous environment.

Bade (2010) proposed an agent-based middleware for sensor networks and RFID systems. This middleware will meet the challenges for having a robust, adaptable and flexible middleware, which is moreover easily extensible to cope with expected re-engineerings and changes while maintaining a clear and elaborate design.

Pitzek (2010) focused the ontologies as the representation of the domain for informational purposes, i.e., as a conceptual domain model, and putting it into context with other domains, such as communication capable devices and automatic identification methods. Basic ontologies for communication protocol, antenna, device, memory, the RFID tag are suggested. Lei (2010) addressed the theory and processes of the coordination mechanism used in RFID network with Multi-Agent System (MAS) framework. It firstly defines an ontology prototype in RFID field; and then gives the basic principle of the mechanism which is used to deal with the collaboration and conflict among agents. This research is a preliminary groundwork for the implementation of ontology based RFID coordination mechanism.

Applications of RFID in different areas are reviewed. Sorensen et al. (2010) represented the case in construction. The paper reviews existing ontologies relevant in relation to creating such a digital link between virtual models and the physical components in the construction and operation phase. The ontologies are reviewed from an ontology consumer system point of view. The ontologies are categorized according to their applicability to specification of technical services, resources, organizational relations, business processes, and overall frameworks for ontology descriptions and their relations. It is concluded, that the technical service and resource ontologies are applicable and that the meta-organizational, and business process ontologies need further development and industrial maturity to be applicable for use in system development. Vrba et al. (2008) presented industrial applications of RFID using agent-based control. They proposed an architecture integrating the RFID technology with the agent-based industrial control solutions by proposing special RFID agents as mediators between physical RFID readers and other agents. In contrast to the physical utilization of RFID technology in warehouse management and supply chain applications, the paper presented architecture for RFID integration at the factory floor level for manufacturing control purposes. Bratukhin and Treytl (2006) addressed the suitability of RFID in agent-based control applications to gain full control over all products and resources at shop floor level. Additionally, an implementation concept for RFID-based product control developed in the EU project PABADIS'PROMISE is presented; the project aims at installing a completely distributed agent-based production control system with full information control.

The purpose of this research is to propose an agent-based control scheme to improve the performance of RFID. The factors which influence RFID performance are RFID device, types of material, object type and environment. In order to integrate these factors to improve the performance, ontology is proposed. Thus, ontology in the RFID domain is proposed and agent-based control scheme is proposed. Specifically, Object ontology, RFID system ontology and overall system ontology are proposed. Rules for ontology are proposed and adopted for reasoning. Section 2 describes the problem domain with overall system structure. Section 3 proposes RFID ontology detail and its applications. Section 4 represents the agent-based control structure. Section 5 summarizes the conclusions and discussions.

2. PROBLEM DOMAIN DESCRIPTION

The system for RFID recognition is composed of object with tag, RFID device, antenna, middleware, legacy system, and ontology. The object is composed of object itself and packaging. The object itself can be plastic, metal, wood, water, snack, books, shoes, etc. According to the types of material, the recognition performance change. Also, the performance is influenced by the packing type. The packing can be paper, plastic, metal or wood. The identification of RFID tag is influenced by the packaging type and object itself.

Environmental factors with the RFID identification are tag type, reader type, reader power level, conveyor speed and location of tag. The location of tag can be front, side and top of the packing box. According to the location, the recognition efficiency differs significantly. The basic function of RFID middleware is monitoring, filtering, analysis and report of EPC data. RFID middleware convert EPC into external applications such as RFID agents, remote GUI, higher level legacy information systems like MES, ERP, SCM or CRM systems.

RDID ontology is related with other systems related with the system. With the RFID tag itself, communication protocol, medium access control, transmission medium, transmission method, duplexing and encoding styles are important factors for the ontology. For the communication protocols such as transmission medium (air, solid), method (light, infrared, laser, sound, radio wave), air and radio wave are chosen for medium and method respectively.

The structure of RFID ontology system is shown in Figure 1. From the supplier, goods with RFID tags are shipped to the warehouse. The goods can be composed of different material and packaging types. With the environmental factors, tag type, conveyor speed, tag location can be considered. Considering RFID device and environmental factors, all relevant data are organized and it is stored as ontology. Between suppliers and warehouse, database of inventory status is shared on real-time for automatic order and supply fulfillment.

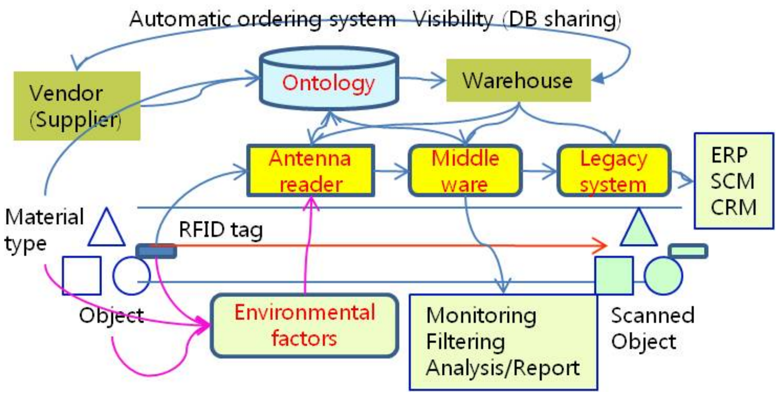


Figure 1. RFID recognition system architecture

3. RFID ONTOLOGY

The recognition rate of RFID differs according to product type and packaging material. In the past when the RFID has been adopted in the warehouse area for the first time, the industry expectation has been so high, but the actual adoption rate is still low level at present because of the tag recognition rate problem. To tackle this problem, though intensive research has been performed about RFID hardware, the result is not so promising because of cost increase related to the RFID tag and devices. Thus, increasing performance of RFID by software-based approach using intelligent method is important.

Jung (2009) reported an intensive survey about RFID performance according to packaging material. He proposed optimal tag location, angle of antenna, conveyor speed, moving speed of forklift for better identification. The research shows that recognition ratio is high when angle between tag and antenna makes 90 degrees in the conveyor portal. For the case of door portal, recognition ratio is optimal when angle between tag and antenna makes 30 degrees. With the location of tag attachment, front location represents optimal result. In this research, though intensive experiment has been performed on the various types of product and packaging material, some issues are still unsolved such as inner and outer packaging material, and interference between product and packaging, etc.

Griffin et al.(2005) studied various object intensively using 915MHz UHF tag. A convenient way to quantify the reduction in tag performance due to material attachment is as a decrease in the RF tag antenna gain, or gain penalty. A gain penalty measurement is a good figure of merit because it combines the effects of a decrease in the radiation efficiency, detuning, and antenna pattern distortion caused by material effects into a single number. They measured loss by wave length in the 3m distance, and investigated power to reach the tag by link budget considering free space loss. The equation (1) describes the amount of power received by the RF tag to operate the tag circuitry (assuming the antennas have an impedance and polarization match):

$$P_{tag} = P_{reader-tx} - L_{sys} + G_{reader-tx} + G_{tag_{-ideal}} - \boldsymbol{Penalty} - 20\log_{10}\left(\frac{4\pi}{\lambda}\right) - 20\log_{10}(d) \quad (1)$$

where Ptag (dB) is the power received by the RF tag, Greader−tx (dB) is the gain of the reader transmit antenna, Lsys (dB) represents the system losses in both the tag and reader, Preader−tx (dB) is the power input to the reader transmit antenna, Gtag−ideal (dB) is the RF tag antenna gain in free space, 20 log10 _4πλ _ is a frequency dependent loss, and 20 log10 (d) is the free space path loss referenced to 1 m.

Using this research, we can forecast objects whether it is feasible to recognize or not by RF. If we are able to know the penalty value of each object in advance, antenna distance can be calculated to recognize the tag attached on the product.

Thus, by saving the penalty value in the knowledgebase and sharing them, the RFID recognition rate can be improved. Especially, creating ontology for the knowledgebase helps to revise and maintain the penalty database by sharing the knowledge. The penalty value which is not known previously can be forecasted using class hierarchy from existing object. The hidden penalty can be searched which is not shown because of different name.

In this research, ontology structure for the object on which a tag is attached is proposed as Figure 2. Object implies a Class. A tag is attached on the material Class. Object has a property with content, inner package and outer package. Each has one individual among the Material Class. Material Class represents a hierarchy of material for RFID tag. Using the WordNet function in the ontology, synonym and homonym can be defined additionally. This research focuses task ontology and agent-based architecture in order to improve RFID performance. Detail about domain ontology is dealt in detail.

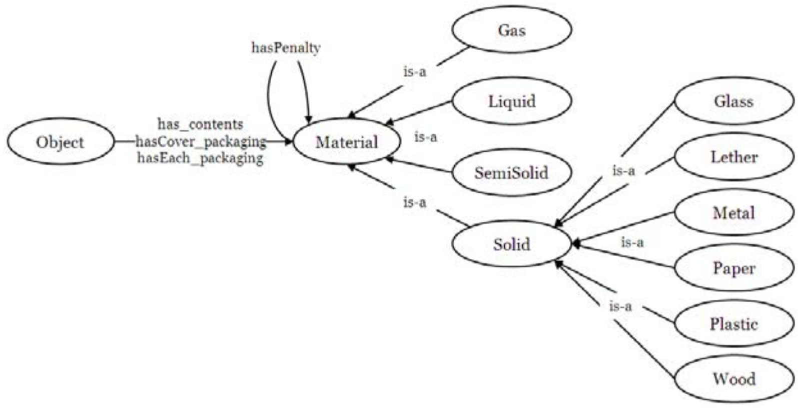


Figure 2. Object Ontology

RFID system ontology is shown in Figure 3. RFID system ontology is a concept which has Hardware, Infra and Software as a lower Class. Hardware Class specifies a hardware system required for RFID system operation. Software Class is related with the middleware. This research focuses more on hardware, and Class for software can be adopted by importing/merging from other existing result. Infra Class is a basic part for RFID system, which has Power and RF as a sub-Class. As Antenna and RF tag should have the same frequency to be recognizable, frequency of other object can be reasoned when one object is determined. Necessary and Sufficient conditions can be defined in Each Class. For example, memory-less tag is read only where there is no memory and only passive type is available (hasPower has Passive). Among the sub-Class of RF Class, MicroWave Class is Active Type and SemiActive type (hasPower has Active; or . hasPower has SemiActive). Through the definition of Necessary & Sufficient condition for each Class, the concept of each Class can be defined explicitly, and this kind of Class satisfying this condition can be reasoned.

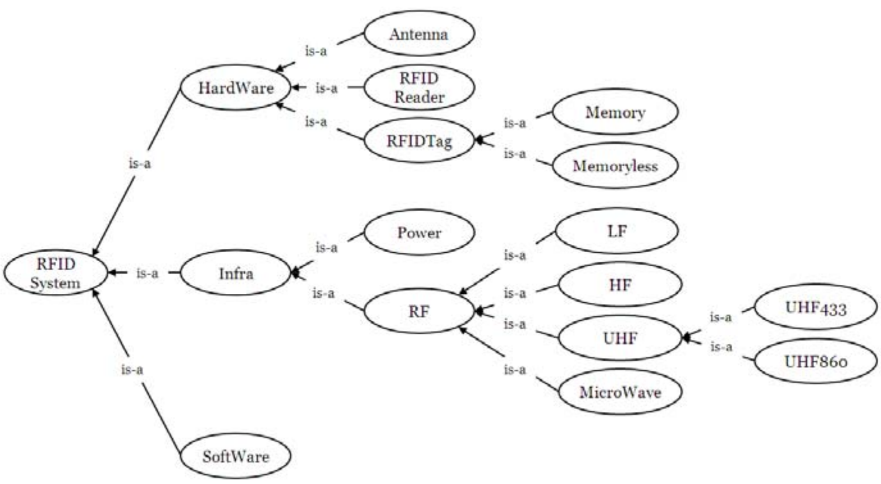


Figure 3. RFID System Ontology

This study proposes "Concept analysis expression" as an intermediary form for collecting expertise. The form is written as {**Concept**: $property_1$, $property_2$, …, $property_n$}. The ontology property contains "asserted property" and "inferred property". The from is described in Table 1.

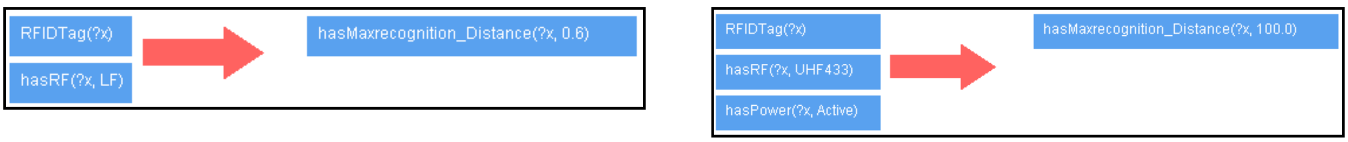
SWRL (Semantic Web Rule Language) is used to reason the inferred property. SWRL provides SWRLM Namespace to calculate a mathematical function. Based on the frequency of RFID Antenna and Tag, Frequency Loss ($20\log_{10}(4\pi/\lambda)$) can be inferred using SWRLM. Also, based on Antenna location, Free Space Loss ($20\log_{10}(d)$) can be inferred. Thus, based on the RFID environmental information, it is diagnosed in advance whether an object with a specified Penalty value can be recognized or not. If the object can't be identified, maximum distance which is identifiable can be calculated by reasoning. We adopted "Rule analysis form" to represent the stepwise inference process. Each step is described using a natural language that can be further translated into a formal rule representation. The form is written as "{**an inferred property**: $Step_1$; $Step_2$; …; $Step_n$}". For example, the solution for finding link budget of a specific object and frequency, distance which is RFID tag and RFID antenna can be given as the following.

{**hasForward_LinkBudget**:
Starting from an object(x) and envirionment(y);
Based on an object(x) to find its penalty(z);
Based on an object(x) to search its RFID Tag(a);
Based on Tag(a) to find its Frequency(b);
Based on Frequency(b) to calculate its Frequency_Dependent_Loss(c);
Based on envirionment(y) to find Antenna_Distance(d);
Based on Antenna_Distance(d) to calculate its Free_Space_Loss(e);
Based on penalty(z) and Frequency_Dependent_Loss(c) and Free_Space_Loss(e) to calculate its Tag voltage(f);
Finally, store voltage(f) "hasForward_LinkBudget" property}

Table 1 Detailed property list of "asserted property" and "inferred property"

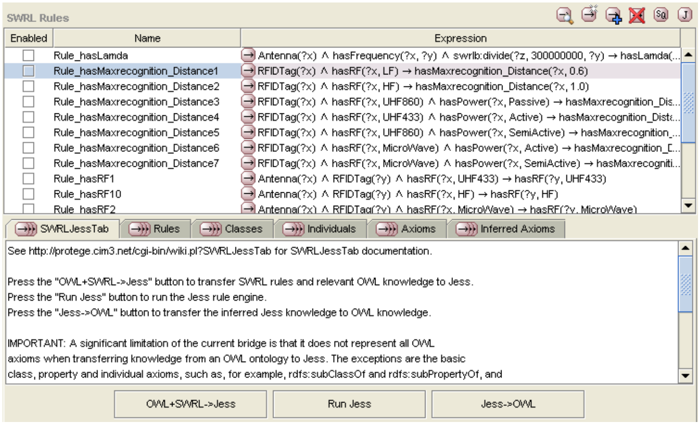
Concept	Property types	Property names
Antenna	Asserted property	hasFrequency, hasOutput, hasReceive_AntennaGain, hasRF, hasTransmit_AntennaGain
	Inferred property	hasLamda
RFIDReader	Asserted property	hasReaderRF_Output, hasRF
	Inferred property	
RFIDTag	Asserted property	has Arrival_Power_dB, hasPower, hasRF, hasSensitivity, hasTag_ArrivalPower, hasTagSensitivity
	Inferred property	hasMaxrecognition_Distance, hasForward_LinkBudget
RF	Asserted property	hasLowFrequency, hasMaxFrequency, hasPower
	Inferred property	hasFrequency_Dependent_Loss

Figure 4 shows reasoning process using SWRL for inferred property. Figure 4-(a) represents the maximum recognition distance is 0.6m if the frequency of RFID tag is LF (Low Frequency). Figure 4-(b) represents the maximum recognition distance is 100m if the frequency of RFID tag is UHF433 and Active Type. In the (a), there is no condition such as hasPower(?x, Active/Passive/SemiActive). As we have defined LF Tag has only Passive Type in the Necessary & Sufficient Condition of LF Class, we need not the condition here.



(a) Maximum recognition distance of LF Tag

(b) Maximum recognition distance of of UHF433, Active



(c) SWRL reasoning rule

Figure 4. Reasoning of inferred property using SWRL

After SWRL-based reasoning rule is generated, it is represented using Axiom Tab of the Protégé Plugin in Figure 4. When the rule is converted into IF THEN structure, it is shown as the following.

IF
"x" IS A RFIDTag
AND "x" HAS RF "LF"
THEN
"x" HAS Maxrecognition Distance "0.6"

Figure 4-(c) represents reasoning rule using SWRL. The reasoning can be performed using JESS (Java Expert System Shell), and the reasoning result can be stored in the inferred property. Also, the inferred property can be used as a parameter for a new inferred rule. In order to calculate Link Budget, the parameter of wave length (λ) is required.

Antenna(?x) $\wedge$ hasFrequency(?x, ?y) $\wedge$ swrlb:divide(?z, 300000000, ?y) $\rightarrow$ hasLamda(?x, ?z)

The above described reasoning rule is an equation to calculate wave length (λ) from the frequency. The inferred wave length is stored in the "hasLamda" as an inferred property. The result is shown in Figure 5.

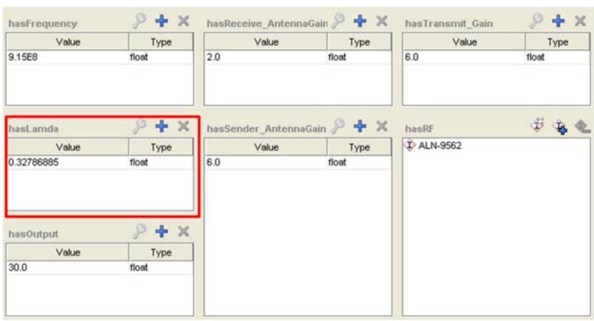


Figure 5. Infered property stored by the SWRL inference rule

The ontology structure developed in this research is shown in Figure 6.

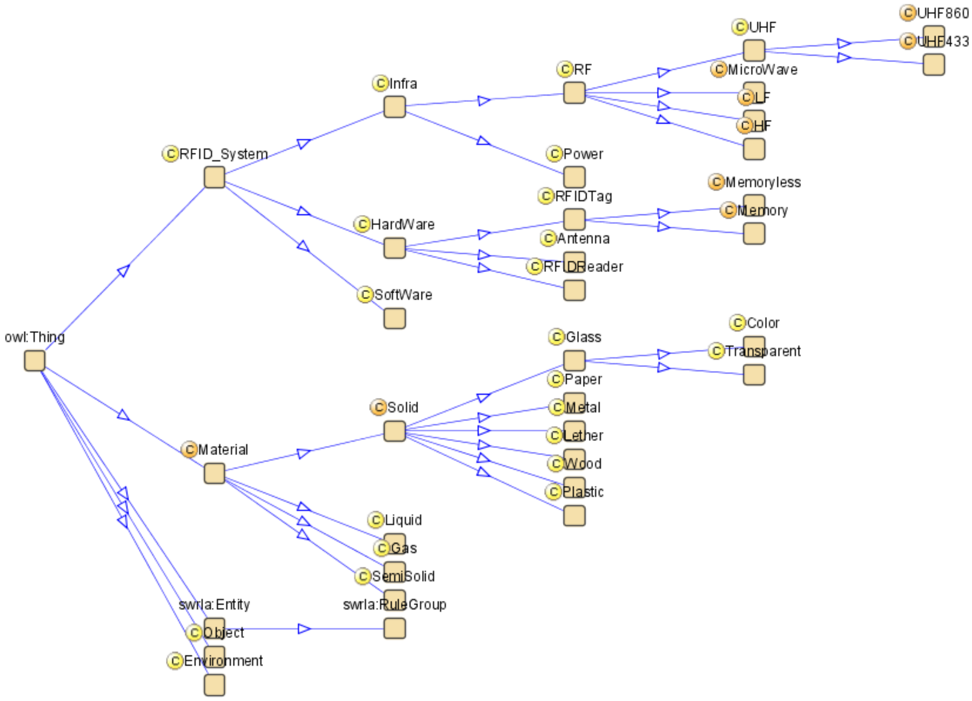


Figure 6. Ontology structure

4. AGENT-BASED CONTROL ARCHITECTURE

In Section 3, RFID ontology model is proposed to improve the recognition ratio. If ontology is integrated with an agent, the efficiency of RFID recognition can be improved more. RFID has been used or reviewed for adoption in many logistics area. In the public procurement service of Korean government, RFID tag is compulsory in the procurement logistics from the year of 2011. Vendors supplying goods to government confronted with urgent situation at hand.

We assume that Gain Penalty of RFID tag is known in this model. Actually in the distribution logistics area, as various products are carried in the conveyor or in other transporter, it is not feasible to identify the material type in advance. However, in the procurement or manufacturing area, after the product item is decided first, the RFID scan is performed based on the promised product, the proposed ontology model in this paper can be adopted to improve the RFID recognition rate.

Figure 6 shows the conceptual model for improving RFID performance by integrating ontology and agent. In the logistics of procurement or manufacturing case, we have a prior information about the product to be inspected or assembled. When simple information about the product is input to the system, agent retrieves Gain Penalty about the product into the RFID ontology model. The ontology model performs searching or reasoning process for the product, and returns data such as Antenna angle and distance which may result in the best recognition rate. Based on the reasoning result from the ontology, the agent performs an action. It will adjust the height of Antenna in the conveyor portal, the loading height in the monitor of forklift truck or the height of forklift when passing the gate. This process helps the recognition rate of RFID tag in varying conditions.

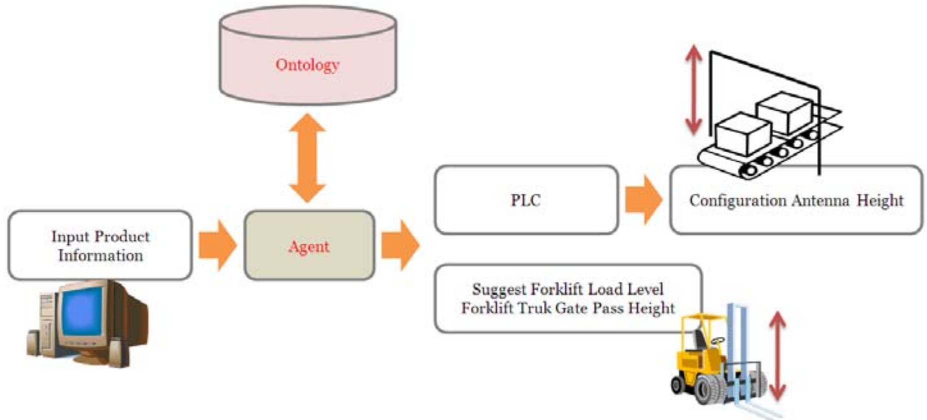


Figure 6. Agent-Based Control Architecture to improve the performance of RFID

5. CONCLUSIONS

In the research, factors affecting the RFID performance is reviewed and analyzed. RFID device, object, material and environmental factors are related with RFID recognition. Ontology architecture is proposed and developed using Protégé 3.0 version. In order to further improve recognition rate, agent based control mechanism is proposed in the conceptual level.

In this research, an assumption is made for the type of product in the incoming goods. The assumption can be reasonable in the procurement and manufacturing areas because the legacy system has prior information for the incoming material. However, in order to be robust, the assumption should be removed and generic model need to be developed to be utilized in the ubiquitous environment.

ACKNOWLEDGEMENTS

This research was finically supported by the Ministry of Education, Science Technology (MEST) and Korea Institute for Advancement of Technology (KIAT) through the Human Resource Training Project for Regional Innovation.

6. REFERENCES

Bade, D. (2009). Towards an Extensible Agent-based Middleware for Sensor Networks and RFID Systems. *Third International Workshop on Agent Technology for Sensor Networks (ATSN-09),* Budapest, Hungary, 1-8.

Bratukhin, A. and Treytl, A. (2006). Applicability of RFID and Agent-Based Control for Product Identification in Distributed Production, http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=04178292, 1198-1205.

Cho, S., Nam, K. and Park, J. (2007) Modeling RFID system in supply chain on manufacturing operation strategy, *Fall Conference of Korean Industrial Engineering*, KunKook University, Seoul, Korea, November 3, 1-10.

Griffin, J, Durgin, G., Haldi, A. and Kipplen, B. (2005). Radio link budgets for 915 MHz RFID antennas placed on various objects, *Texas Wireless Symposium*, 22-26.

Jung, S. (2009) Packaging guideline for applying RFID/USN, *Fall Conference of Korean Industrial Engineering*, Kyunghee University, Suwon, Kyung-Gee-Do, October 9, 1-77.

Lei, T. (2010). Ontology based coordination in RFID network, http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=04680536, 1-4.

Pitzek, S. (2010). An ontology for the RFID domain, http://techscreen.at/Ontologies/RFIDOntology0903/RFIDOntology-Report.pdf, 1-11.

Sorensen, K. Christiansson, P. and Svidt, K. (2010). Ontologies to Support RFID-Based Link between Virtual Models and Construction Components. *Computer-aided Civil and Infrastructure Engineering, 25*: 285-302.

Vrba, P. Macurek, F. and Marik, V. (2008) Using radio frequency identification in agent-based control systems for industrial applications. *Engineering Applications of Artificial Intelligence, 21:* 331-342.

Session F3: Seaport and Transportation 6

·Day2: Sep. 16 (Thu.)

·Time: 10:40 - 12:00

·Chair: Hans-Otto Günther

·Room: Iris, 4F

PORT CHOICE BEHAVIOR OF LINERS UNDER COST AND DEMAND UNCERTAINTY - FROM THE PERSPECTIVE OF NODE GAMES

Xiaoning Shi[1*] and Stefan Voß[1]

[1] University of Hamburg
Germany
e-mail: shi@econ.uni-hamburg.de
stefan.voss@uni-hamburg.de

Abstract: Besides cooperation and competition among liners in the global shipping market, a liner might build up vertical collaboration together with selected port operators in terms of performing port choice behavior. On one hand, the liner makes decision taking into account cost uncertainty caused by different cargo handling systems and dynamic transportation demand derived from on-going international trade. On the other hand, port operators, in particular those adjacent ones providing similar service, try hard to become attractive nodes so that they can be chosen by the liner as ports-of-call.

In this paper, we start by analyzing port choice behavior of liners given static transportation demand. In the static phase, by improving cargo handling efficiency as well as offering lower prices, port operators can significantly affect liners' port choice behavior. In the next phase, dynamic demand is to be considered in the model, and more interactive node games are constructed to simulate procedures of continuing port choice behavior of related liners. In addition, sensitivity analysis is performed regarding changing cost as well as uncertain demand. Finally, suggestions to liners and port operators are summarized.

SHORT-TERM ROUTING AND TRANSPORTATION PLANNING IN SPOKE NETWORK IN THE BLACK SEA REGION

Olcay Polat[1], Osman Kulak[2], and Hans-Otto Guenther[1*]

[1] Technische Universität Berlin
Germany
e-mail: olcaay@gmail.com
Hans-Otto.Guenther@tu-berlin.de

[2] Pamukkale University
Turkey
e-mail: okulak@pau.edu.tr

Abstract: Container ship liners plan their network design based on long-term transportation demand and fixed routes and schedules for hub-to-hub services within hub-and-spoke systems. In the short run, however, hub-to-spoke transportation quantities, routes and schedules depend on the supplies and demands at the spoke ports. For long-term considerations these figures are highly uncertain due to varying transportation demand. In our investigation, various spoke network design strategies are analyzed by use of a simulation model into which heuristic algorithms for short-term transportation planning are embedded. The proposed model implementation has been tested using a case study from the Black Sea region. It is shown that the short-term routing and transportation planning problem has a major impact on the design of spoke networks in regional maritime container shipment.

THE CAPACITY EXPANSION PROBLEM IN CONTAINER TERMINALS

Anh-Dung Do Ngoc[1] and Il-Kyeong Moon[1]*

[1]Department of Industrial Engineering
Pusan National University
Busan 609 – 735
Korea
E-mail: dnadung@pusan.ac.kr
ikmoon@pusan.ac.kr

1. INTRODUCTION

Recently, container shipping has become more and more popular, which means the demand for full container transportation is rapidly increasing. That is a challenge to shipping companies. On one hand, they are able to yield more profit by providing empty containers and transporting full containers to required destinations. On the other hand, there is an imbalance among the regions in empty container requirement. The ports in US or Europe have a surplus of empty containers, while the ones in East Asia have deficits. As a result, shipping companies are under the influence of this imbalance. In East Asia, the shipping companies need to stock more empty containers to meet the requirements, while paying for holding empty containers at depots in the surplus regions. Empty container positioning is one of the most effective ways to overcome this problem. Empty containers from surplus regions will be transported to deficit ones to reduce the imbalance. There have been several studies considering empty container positioning and they pointed out that the cost can be saved. Unfortunately, we cannot position as many empty containers as possible since there is a limitation in terms of the depot capacity. Moreover, the demand is usually fluctuating. Empty containers should be kept at depot in the low demand period and used when the demand becomes high. A shipping company will be more competitive if the depot capacity is large. Depot capacity expansion plays a significant role in the activities of shipping companies.

Capacity expansion is not a new issue. Many studies have been conducted to consider this problem. However, there have been no studies on the capacity expansion of container depots. In this research, we provide a decision tool to plan the expansion of depot capacity. If the capacity is expanded too early, it will be wasteful and we need to pay for the unused capacity. The budget for expansion could be invested in other projects to yield more profit. On the other hand, it is very hard and costly for the company to satisfy the demand for empty containers if the depot capacity is expanded too late. Another situation is that the company needs to lease empty containers, which usually entails higher cost. Therefore, we need a plan to determine when we have to expand the capacity and how large it should be. An alternative way to enlarge the depot capacity is to lease storage space from other depots. In each period, the storage space can be leased with adequate amounts in several periods and be returned later.

As mentioned above, depot capacity expansion affects the number of empty containers to be held in the company's own depot or repositioned from other ports. Hence, in this research, we will consider not only capacity expansion planning, but also plans for positioning and leasing. The mathematical model is to minimize the total relevant costs consisting of the capacity expansion cost, storage space leasing cost, inventory holding cost, container leasing cost, and positioning cost. There is a limitation on the capacity expansion and leased storage space. The decision variables involve the amount of expanded capacity, the amount of leased storage space, the number of leased containers, and the number of positioned containers.

2. PROBLEM DESCRIPTION

2.1. Literature Review

Recently, many studies have been carried out concerning empty container positioning. Crainic *et al.* (1989) considered a multi-commodity location problem with balancing requirements. They proposed models for multi-commodity capacitated location problems (MCLB) with an inter-depot balancing requirement. Many studies have been carried out in an attempt to solve the MCLB problem. Crainic and Delorme (1993) developed dual-ascent procedures for the proposed model. Crainic *et al.* (1993) solved the problem by using a Tabu search procedure. Gendron and Crainic (1997) presented a parallel branch and bound approach, which is based on the dual-ascent procedure previously proposed by Crainic and Delorme (1993). Gendron *et al.* (2003) also solved the problem using a Tabu search procedure, but used the slope scaling method to find the starting solution. Li *et al.* (2004) studied the management of empty containers in a port with stochastic demand. Their analysis was based on a multistage inventory problem and Markov decision processes with discrete time. They focused on the optimization of the pair-critical policy, (U, D). In this policy, if the number of empty containers at a port is less than U, empty containers are imported up to the amount of U; if it is more than D, empty containers are exported down to the amount of D. Li *et al.* (2007) extended the problem for multipleports. Shen and Khoong (1995) proposed a decision support system (DSS) for empty container distribution planning. The DSS is based on network optimization models. In the network, they considered the leasing-in, off-leasing, positioning-in, and positioning-out at a port. The problem was decomposed into three levels, namely, terminal (port) planning, intra-regional planning, and inter-regional planning. They considered a single type of container.

2.2. Problem Definition

The purpose of this paper is to investigate the expansion of depot capacity. This research will help the decision maker to decide when the depot capacity should be expanded as well as the amount of expanded capacity. In this research, we also consider the positioning and leasing of empty containers. The problem is categorized as lying between tactical and strategic planning. There are several ports/depots and each of them has demand for empty containers in each period. When a port has a shortage of empty containers, leasing will be carried out to satisfy demand. On the other hand, if the port has a surplus of empty containers, these containers will be either repositioned to other depots that have shortage of empty containers or stored in inventory in reserve for future demand. There are two alternatives for expanding the depot capacity: expanding the current depot or leasing storage space. The amount of expanded capacity is limited but the number of leased containers is not.

2.3. Mathematical Model

The objective function is to minimize the total relevant cost. The total cost consists of the cost of expanding depot capacity, cost of leasing storage space, cost of positioning empty containers, leasing cost, and holding cost. The constraints consist of inventory balance equations, limitation on inventory capacity, limitation on the amount of expanded capacity, and limitation on the amount of leased storage-space. As in many studies on empty container movement, especially those of Crainic *et al.* (1989, 1993) and Crainic and Delorme (1993), the variables that relate to the flow of empty containers are considered continuous variables in this study. The values of these variables are large and can be rounded to the nearest integer.

3. THE SOLUTION ALGORITHM

The mathematical model is a Mixed Integer Programming, and therefore, when the size of the problem increases, optimization software cannot solve it in a reasonable period of time. For that reason, we apply Largrangian Relaxation (LR) method to solve the problem with short computational time and an acceptable solution. The LR problem is formed by relaxing the inventory capacity constraint from the original problem and it (i.e. LR problem) can be decomposed into three sub-problems that relate to the activities of expanding capacity, leasing storage space, and positioning empty containers. These sub-problems can be easily and quickly solved by a proposed heuristic algorithm. The optimal value of the objective of LR problem will be a lower bound on that of the original problem and it (i.e. the optimal value of the objective of LR problem) will be used for evaluating the proposed solution algorithm. The feasible solution is found based on the solutions of inventory levels, the flows of empty containers, and the number of leased containers from the LR problem. A heuristic algorithm is proposed to find the feasible solution.

4. COMPUTATIONAL EXPERIMENTS

All the experiments are conducted on a computer with an Intel Core 2 Quad 2.4 GHz and 3.24 GB of RAM. LINGO was used to solve the empty container positioning problem. About 30 problem instances that have different relevant costs, maximum expanded capacities, demands, and other parameters are generated in order to evaluate the performance of the proposed algorithm. These problem instances are categorized in small size, medium size, and large size. The small size problems have 5 to 10 ports, 4 to 6 container types, and 30 to 50 periods. The medium size problems have 11 to 20 ports, 6 to 8 container types, and 50 to 75 periods. The large size problems have 11 to 20 ports, 8 to 10 container types, and 75 to 100 periods. Table 1 presents the result of the experiments.

Table 1. Results of the experiments

Problem Instance	MIP		Lagrangian based algorithm			% Penalty
	Obj. Value	Comp. Time (second)	Obj. Value	Lower Bound	Comp. Time (second	
1	43,194,933	0.88	43,218,808	43,154,601	0.45	0.06%
2	52,856,104	1.05	53,063,513	52,856,104	0.53	0.39%
3	96,082,710	1.41	97,380,452	96,082,710	0.76	1.35%
4	40,845,607	0.66	41,033,128	40,845,607	0.38	0.46%
5	144,904,752	2.52	145,457,779	144,904,752	1.14	0.38%
6	37,223,996	0.89	38,513,462	37,223,996	0.50	3.46%
7	70,857,347	1.02	71,628,511	70,857,347	0.52	1.09%
8	169,171,301	2.31	169,520,179	169,171,301	1.20	0.21%
9	48,420,821	1.13	49,036,631	48,420,821	0.61	1.27%
10	87,530,512	1.69	87,538,216	87,530,512	0.84	0.01%
11	442,382,553	20.28	443,018,610	442,382,553	6.10	0.14%
12	281,808,360	13.47	283,083,705	281,808,360	4.45	0.45%
13	756,934,208	52.11	757,421,433	756,934,208	11.62	0.06%
14	508,866,914	36.53	516,318,625	508,866,914	9.16	1.46%
15	1,011,116,279	63.75	1,018,409,781	1,011,116,279	15.00	0.72%
16	404,814,572	22.05	405,503,520	404,814,572	5.86	0.17%
17	361,452,650	16.23	362,556,652	361,452,650	5.08	0.31%
18	315,560,152	13.25	315,726,649	315,560,152	4.31	0.05%
19	454,018,478	30.89	454,018,478	454,018,478	8.41	0.00%
20	765,818,852	56.42	766,254,414	765,818,852	13.33	0.06%
21	2,371,634,377	522.80	2,372,819,917	2,371,634,377	67.23	0.05%
22	1,685,537,098	423.77	1,689,039,500	1,685,537,098	42.35	0.21%
23	-	-	3,016,387,571	3,015,983,049	97.33	0.01%
24	-	-	2,478,193,528	2,475,390,529	87.03	0.11%
25	1,486,644,563	296.28	1,488,192,517	1,486,644,563	28.18	0.10%
26	1,217,690,558	245.22	1,218,345,115	1,217,690,558	25.56	0.05%
27	-	-	2,314,683,724	2,314,178,688	76.98	0.02%
28	1,972,930,283	493.63	2,002,020,808	1,972,930,283	58.35	1.47%
29	2,203,155,300	586.45	2,211,802,376	2,203,155,300	67.99	0.39%
30	1,753,921,241	674.00	1,755,812,518	1,753,921,241	57.62	0.11%

In Table 1, the cells containing a "-" symbol indicate that the corresponding problems cannot be solved within a day and results cannot be obtained. For each problem instance, the computational time for the LR-based algorithm is the average computational time over ten runs. Similarly, the percentage penalty for the LR-based algorithm is calculated by using the average objective value of ten runs. In those problems where we cannot obtain the optimal solution, the percentage penalty is calculated by using the lower bound. It can be seen that the LR-based algorithm can solve large-sized problems very quickly. Moreover, the percentage difference between the LR-based algorithm and MIP does not exceed 4%. These features indicate good performance of our proposed algorithm.

5. CONCLUSION

This study considers the capacity expansion problem for container terminals or depots. A mathematical model is built to minimize the total relevant cost for multiple-ports, multiple-commodities, and multiple-periods. A solution algorithm based on Lagrangian Relaxation is proposed. This study can be used to support decision making between short-term and long-term planning horizons. The experiments are conducted for up to 30 ports, 10 container types, and 100 periods. From the experiments, it can be seen that the proposed algorithm can solve the problem very quickly (in approximately one minute) with less than 5% error. These are persuasive indications of the effectiveness of the proposed algorithm.

6. REFERENCES

Berman, O. and Ganz, Z. (1994). The Capacity Expansion Problem in The Service Industry. *Computers Operations Research, 21:* 557–572.

Crainic, T. G., Dejax, P. and Delorme, L. (1989). Models for Multimode Multicommodity Location Problems with Interdepot Balancing Requirements. *Annals of Operations Research, 18:* 279 – 302.

Crainic, T. G., and Delorme, L. (1993). Dual-Ascent Procedures for Multicommodity Location-Allocation Problems with Balancing Requirements. *Transportation Science, 27:* 90 – 101.

Crainic, T. G., Gendreau, M., Soriano, P. and Toulouse, M. (1993). A Tabu Search Procedure for Multicommodity Location/Allocation with Balancing Requirements. *Annals of Operations Research, 41:* 359 – 383.

Fisher, M. L. (1985). An Application Oriented Guide to Lagrangian Relaxation. *INTERFACES, 15*: 10–21.

Gendron, B. and Crainic, T. G. (1997). A Parallel Branch-and-Bound Algorithm for Multicommodity Location with Balancing Requirements. *Computers and Operations Research, 24:* 829 – 847.

Gendron, B., Potvin, J. Y. and Soriano, P. (2003). A Tabu Search with Slope Scaling for the Multicommodity Capacitated Location Problem with Balancing Requirements. *Annals of Operations Research, 122:* 193 – 217.

Guignard, M. (2003). Lagrangean Relaxation. *Top, 11:* 151 – 228.

Li, J. A., Liu, K., Leung, S. C. H. and Lai, K. K. (2004). Empty Container Management in A Port with Long-run Average Criterion. *Mathematical and Computer Modelling, 40:* 85 – 100.

Li, J. A., Leung, S. C. H., Wu, Y. and Liu, K. (2007). Allocation of Empty Containers between Multi-ports. *European Journal of Operational Research, 182:* 400 – 412.

Luss, H. (1982). Operation Research and Capacity Expansion Problem: A Survey. *Operations Research, 30:* 907–947.

Min, H. and Ko, H. J. (2008). The Dynamic Design of A Reverse Logistics Network from The Perspective of Third-Party Logistic Service Providers. *International. Journal of Production Economics, 113:* 176–192.

Moon, I. K., Do Ngoc, A. D. and Hur, Y. S. (2010). Positioning Empty Container among Multiple Ports with Leasing and Purchasing Consideration. *OR Spectrum, 32:* 765 – 786.

Shen, W. S. and Khoong, C. M. (1995). A DSS for Empty Container Distribution Planning. *Decision Support Systems, 15:* 75 – 82.

A QUANTITATIVE APPROACH FOR SUPPORTING THE ADAPTATION OF SEAPORTS TO MODERN SUPPLY CHAIN CONCEPTS

Frank Schwartz[1] and Stefan Voß[1*]

[1] University of Hamburg
Germany
e-mail: fs@econ.uni-hamburg.de
stefan.voss@uni-hamburg.de

Abstract: One important approach in supply chain management is to spread out the production process to several locations within the supply chain. This concept is called postponement and provides a better adaptation of standardized manufacturing items to regional needs. The realization of the postponement concept requires the implementation of several value added logistics services in the affected sites. Ideal locations for implementing the required facilities are seaports because they represent central transition points for larger regions.

Considerable problems for the implementation of postponement activities and the associated value added logistics services in seaports are spatial restrictions which hinder the construction of the required additional buildings and storage areas. A solution to this problem is establishing dry ports which provide the possibility to locate the required plants for postponement activities in the neighborhood of seaports. Important aspects in this context are decisions if the value added logistics services ought to be implemented in an existing sea port anymore or in a dry port. Further decisions concern potential infrastructure investments to ensure the designated flow of goods between seaports and dry ports.

In the presented paper a mixed integer linear programming (MILP) model is developed that permits the support of decisions coming up when an expansion of postponement activities and the realisation of a dry port concept are intended. Several planning scenarios are examined and the solvability of the model is demonstrated by computing some test instances that mimic realistic supply chain considerations.

Session F4: Industrial Session 2

· **Day2: Sep. 16 (Thu.)**

· **Time: 10:50 - 12:20**

· **Chair: Kwang Ryel Ryu**

· **Room: Grand Ballroom, 5F**

AUTOMATION AT PATRICK CONTAINER TERMINALS: THE AUTOSTRAD™ SOLUTION

Daniel Pagac

Technical Manager
Patrick Technology and Systems
Australia
Email: d.pagac@patrick.com.au

Abstract: Asciano Ltd is one of Australia's largest infrastructure operators operating rail assets, ports assets and associated services. Its annual revenues exceeded $2.8 billion and the company is well positioned to benefit from the continued growth in Australian trade. Its subsidiary, Patrick Container Ports (PCP), owns and operates leading container terminals in Brisbane, Sydney, Melbourne and Fremantle. PCP's Brisbane terminal at Fisherman Islands is a fully automated medium density operation using straddle carriers (AutoStrads®). PCP has developed this technology to a stage where it now provides a clear global leadership position in container terminal automation.

The AutoStrad is based on a free-ranging navigation platform integrated with a straddle carrier control system making it simple and elegant solution for total terminal automation. The sophistication of the control systems meets or exceeds productivity of comparative manual terminals. By pooling of resources the focus is on efficiently solving the transport and logistic problems. This presentation will focus on these two important aspects of a fully automated terminal; the transport problem, essentially the spatiotemporal planning and plan execution of a fleet of AutoStrads and the logistic problem, addressing the operational requirements for the movement of containers through the terminal. References will be made to the broader logistic chain.

Session G1: SCM 6

· **Day2: Sep. 16 (Thu.)**

· **Time: 13:00 - 14:20**

· **Chair: SangHeon Lee**

· **Room: Camellia, 5F**

HEURISTIC APPROACH FOR THE UNCAPACITATED MULTIPLE TRAVELING PURCHASER PROBLEM

Myung-Jin, Choi[1] and Sang-Heon, Lee[1*]

[1]Department of Operations Research
Korea National Defense University
Susaek-dong 205, Eunpyung-gu
Seoul, Korea
e-mail: sanghnlee@hanmail.net

Abstract: The traveling purchaser problem (TPP) is a generalization of the well-known traveling salesman problem (TSP) and has many real-world applications such as purchasing required raw materials for manufacturing factories, scheduling a set of jobs for certain machines, and many others. TPP could be used for inbound logistics optimization in the area of supply chain management. A multiple TPP (MTPP) is to select a subset of markets and route vehicles to the selected markets such that the demand for each product is satisfied and the total purchasing and travel costs are minimized. If a product is available at a given market and its quantity is sufficient to satisfy the demand, this version of the MTPP will be called the uncapacitated multiple traveling purchaser problem (UMTPP). In this case, split purchases are not considered. The purpose of this paper is to introduce the mathematical model for the UMTPP and to propose a heuristic algorithm for the UMTPP. The performance of the proposed algorithm for some generated instances is compared with the optimal solution calculated by CPLEX. By considering the computational complexity, the experimental results show that the proposed algorithm is effective for those instances.

1. INTRODUCTION

The traveling purchaser problem (TPP) as a generalized type of well-known traveling salesman problem (TSP) is defined as follows. Let us consider a set of products to be purchased by a purchaser (vehicle) who is originally at a depot. There is a requirement of units for each different product. Let us also consider a set of markets, each selling some units of a certain number of products. The price of a product depends on the market where it is available. The travel cost between the two markets is also known. The TPP selects a subset of markets and routes a vehicle to the selected markets such that the demand for each product is satisfied and the total purchasing and travel costs are minimized. It is assumed that each product is available in at least one market, no product is available in the depot and the required demand must be purchased. It is also assumed that if a product is available at a given market, its quantity is sufficient to satisfy the demand. This version of the problem will be called the uncapacitated traveling purchaser problem (UTPP). If the availability of a product at a market is smaller than the demand with limited product availability at some markets, then this version of the problem will be called the capacitated traveling purchaser problem (CTPP). It is also assumed that the travel cost between each market is symmetric. This will be called the symmetric traveling purchaser problem (STPP). Having no balance will be the asymmetric traveling purchaser problem (ATPP). The TPP is known to be NP-hard in the strong sense to TSP when each product can be purchased in exactly one market which is not provided by the remaining ones.

Ramesh (1981) proposed the TPP for the first time and developed two algorithms for solving the UTPP. The first one is the lexicographic search, which is an exact algorithm to find out the optimal solution. It's only available, though, for a very small scale problem. The second is the heuristic method, which is called a near neighbor algorithm and which inserts a market into the tour and extends a tour by inserting other markets into then tour repeatedly in a greedy way. Goden et al. (1981) suggested the GSH (generalized savings heuristic) for solving the UTPP. It starts with an initial solution, including the depot and the market selling more products than any other market at the cheapest price. The GSH selects a market based on the savings calculated from the travel and purchase costs and inserts that market into a current tour repeatedly and terminates when no more savings can be made. Ong (1982) proposed the TRH (tour reduction heuristic), which is an improved version of the GSH. It starts with an initial tour consisting of selected markets that collectively carry all the products and repeats the market drop procedure (which drops a market from the current tour for cost reduction). It is terminated when no more cost savings can be made. Voß (1996) presented the meta-heuristic approaches for the UTPP based on a dynamic tabu search and simulated annealing. He proposed two dynamic strategies for the management of the tabu list : the reverse elimination method and the cancellation sequence method. He also introduced two heuristic methods for constructing the initial tour: the ADD-procedure and the DROP-procedure. Singh and Oudheusden (1997) developed a branch-and-bound algorithm for the TPP. The basic idea of this algorithm is that the selection of a subset of markets for the tour and the determination of an optimal tour of these markets are successfully embedded into one. Pearn and Chien (1998) improved the two previous works of Golden et al. (1981) and Ong (1982). They suggested three heuristic methods: the PS-GSH (parameter selection GSH), the TS-GSH (tie selection GSH), and the CAH(commodity adding heuristic). Laporte et al. (2003) researched the CTPP for the first time. They developed the MAH (market adding heuristic) for the UTPP and CTPP. Boctor et al. (2003) proposed three PH (perturbation heuristics) based on a tabu search: UPH1 and UPH2 for the UTPP, and the CPH for the CTPP. These three heuristics comprise seven basic procedures: market drop, market add, market exchange, the TSP heuristic, cheapest insertion, double market drop, and double market exchange. Teeninga and Volgenant (2004) introduced pre-processing and intensification procedures for three previous algorithms, like the GSH, CAH, and TRH. Riera-Ledesma and Salazar-Gonalez (2005a) suggested an LS (local search) algorithm. Riera-Ledesma and Salazar-Gonalez (2005b) also proposed the 2TPP (bi -objective TPP). The objective of the 2TPP is to minimize both the travel cost and the purchase cost simultaneously. Riera-Ledesma and Salazar-Gonalez (2006) introduced an exact algorithm for the ATPP. They discovered an optimal solution for the ATPP by using a branch-and-cut algorithm and studied the conversion from the ATPP to the STPP.

In recent years, Bontoux and Feillet (2008) proposed the DMD-ATA (dynamic multidimensional anamorphic traveling ants algorithm) based on ACO (ant colony optimization) and an LS algorithm. Goldbarg et al. (2008) proposed a TA (transgenetic algorithm) based on horizontal gene transfer (Jain et al., 2003) and endosymbiosis (Margulis, 1992). Angelelli et al. (2009) introduced D-TPP (dynamic TPP), in which the quantity of products at markets is decreased due to the passage of time. For solving this, a few greedy heuristics are suggested: Product-operating, Market-operating, Consumption-operating, and Tradeoff-operating criteria. Mansini and Tocchella (2009) proposed the TPP-B (TPP with budget constraint), in which the travel cost is minimized in an objective fuction and the purchase cost is restricted within a constraint.

The number of purchasers in recent TPP studies is limited to a single purchaser. It could prevent TPP being applied to a real world inbound logistics optimization problem. Therefore, Choi and Lee (2009) suggested the periodic heterogeneous multiple TPP based on the capacitated multiple TPP (CMTPP) for refuse logistics optimization. Choi and Lee (2010), also, proposed the heterogeneous multiple TPP with a budget constraint. However, they didn't suggest a method for solving the MTPP. In this paper, we consider several purchasers and propose the uncapacitated multiple TPP (UMTPP). In this case, the number of the depots is 1, and each purchaser (vehicle) has a purchase (loading) capacity. The UMTPP could be used for inbound logistics optimization in the area of supply chain management and many others. We introduce the mathematical model for the UMTPP and suggest heuristic algorithm for solving this.

2. MATHEMATICAL MODEL

2.1 Brief introduction and assumptions

As a generalization of the TPP, the UMTPP to be reviewed in this section means that several vehicles of a given capacity originally start and end the route at a depot and minimize the total cost of purchasing and travel, while achieving the satisfaction of demand. In order to build a mathematical model of the UMTPP, the following constraints are assumed:

(i) Several vehicles with a given capacity are originally at a depot.
(ii) Each vehicle starts and ends the route at the depot.
(iii) A vehicle visits a market at most once and products purchased in markets along the vehicle's path will be transported by the vehicle to the depot.
(iv) Each product is available in at least one market and no product is available in the depot.
(v) The number of products purchased must be sufficient to satisfy the demand.
(vi) If the demand for products is satisfied then it is not mandatory to visit every market. This is different from the TSP and VRP (vehicle routing problem).
(vii) There is a fixed upper limit capacity for each vehicle and the capacity of all vehicles is identical (homogeneous).

2.2 Formulation of the UMTPP

In this section we present an ILP (integer linear programming) formulation of the UMTPP. This model can be used whether travel costs between each markets is symmetric or asymmetric because we use a three-index type formulation that uses the binary variable x : variable x_{ijk} counts the number of times an arc (i, j) is traversed by vehicle k in an optimal tour.

We summarize notation and decision variables as follow.

<Notation>

m : the number of markets indexed i , $i \in M = \{1,\cdots,m\}$; $i = 1$: depot

M^- : the set of markets without a depot, $M^- = M \setminus \{1\}$

n : the number of products indexed l , $l \in N = \{1,\cdots,n\}$

v : the number of vehicles indexed k , $k \in V = \{1,\cdots,v\}$

b_{il} : the price of product l at market i

c_{ij} : the travel cost between markets i and j ; $j = 1$: depot

d_l : the demand for product l , $d_l = 1$, $\forall l \in N$

q_{il} : the quantity of product l available at market i , $q_{il} \in \{0,1\}$, $\forall i \in M^-$, $\forall l \in N$

C : the capacity of each vehicle

<Decision variables>

$$x_{ijk} = \begin{cases} 1 & \text{If vehicle } k \text{ v isits market } j \text{ immediately after } i, \\ 0 & \text{Otherwise,} \end{cases}$$

$$y_{ikl} = \begin{cases} 1 & \text{if vehicle } k \text{ purch ases product } l \text{ at market } i, \\ 0 & \text{Otherwise,} \end{cases}$$

u_i , u_j : variables for sub-tour elimination .

The formulation of the UMTPP is as follow.

$$\min \sum_{i\in M} \sum_{j\in M} (c_{ij} \sum_{k\in V} x_{ijk}) + \sum_{i\in M^-} \sum_{l\in N} (b_{il} \sum_{k\in V} y_{ikl}) \qquad (1)$$

Subject to

$$\sum_{i\in M} \sum_{k\in V} x_{ijk} \le 1 \qquad \forall j \in M^-, \qquad (2)$$

$$\sum_{j\in M^-} x_{1jk} = 1 \qquad \forall k \in V, \qquad (3)$$

$$\sum_{j\in M^-} x_{j1k} = 1 \qquad \forall k \in V, \qquad (4)$$

$$u_i - u_j + m \sum_{k\in V} x_{ijk} \le m-1 \qquad 2 \le i \ne j \le m, \qquad (5)$$

$$\sum_{i\in M^-} \sum_{k\in V} y_{ikl} = d_l \qquad \forall l \in N, \qquad (6)$$

$$\sum_{j\in M} x_{ijk} - y_{ikl} \ge 0 \qquad \forall i \in M^-, \forall k \in V, \forall l \in N, \qquad (7)$$

$$\sum_{j\in M} x_{jik} - y_{ikl} \ge 0 \qquad \forall i \in M^-, \forall k \in V, \forall l \in N, \qquad (8)$$

$$q_{il} - \sum_{k\in V} y_{ikl} \ge 0 \qquad \forall i \in M^-, \forall l \in N, \qquad (9)$$

$$\sum_{i\in M^-} \sum_{l\in N} y_{ikl} \le C \qquad \forall k \in V, \qquad (10)$$

$$x_{ijk} \in \{0,1\} \qquad \forall i \in M, \forall j \in M, \forall k \in V, \qquad (11)$$

$$y_{ikl} \in \{0,1\} \qquad \forall i \in M^-, \forall k \in V, \forall l \in N, \qquad (12)$$

$$d_l = 1 \qquad \forall l \in N, \qquad (13)$$

$$q_{il} \in \{0,1\} \qquad \forall i \in M^-, \forall l \in N. \qquad (14)$$

Objective function (1) m eans m inimizing th e total sum of t he tra vel cost and purchase cost in t he optimal to ur. Constraint (2) means a ve hicle is allowed t o visit the m arket at m ost once, a nd if the dem ands for all products are satisfied i n som e selected m arkets the ve hicles do not need t o visit the rem aining m arkets. Const raint (3) m eans all vehicles at the depot can start only one time. Constraint (4) means all vehicles m ust return to the depot. Constraint (5) formulates the SECs (subtour elimination const raints), which was proposed by Miller et al. (1960). It's an extended version of Miller et al.'s SECs in the TSP and could be applied in the UMTPP. Constraint (6) means that exactly the number of products required shall be purchased. Constraints (7) and (8) mean that the market should be included in the tour prior to purchasing a product in those markets. Constraint (9) means that a purchaser can purchase a product in the market if the product is in that market. Constraint (10) represents the given capacity of each vehicle. Constraints (11) ~ (14) show the variable's integer conditions.

3. SOLUTION ALGORITHM

In this section we describe the proposed heuristic algorithm. This will be called the vehicle allocation heuristic (VAH). The VAH is constructed with 3 basic heuristics: initial tour constitution, tour improvement, and vehicle allocation based on GSH (Golden et al., 1981) and TRH (Ong, 1982). At the initial tour constitution step, we make a feasible initial tour by using the revised GSH. In GSH, at the beginning of the algorithm, the directed travel costs are considered. However, in the revised GSH, the shortest travel costs are considered at the beginning of the algorithm. At the tour improvement step, we improve that tour by using the revised TRH. In TRH, all the markets in the initial tour are chosen by the decision maker. However, in the revised TRH, markets in the initial tour are chosen by the result of the revised GSH. After the tour improvement step, we apply the 3-opt algorithm (Lin, 1965) to the single tour for further reduction of the travel cost. Finally, at the vehicle allocation step, we allocate the markets in the tour to the vehicles.

At the initial step, we have a depot (v_1), the current tour (α), any market not in α (q), the cost of product l at market i ($b(i,l)$), and the cost of travel from market i to j ($c(i,j)$). We also have the cheapest cost of product l in α ($f(\alpha,l)$), the decrease in the cost of product l ($g(\alpha,q,l)$), and the savings if q is inserted between two adjacent markets i and j in α ($s(i,j,q)$). These are defined as follows:

$$f(\alpha,l) = \min\{b(i,l) \mid i \in \alpha\} \tag{15}$$

$$g(\alpha,q,l) = \max\{f(\alpha,l) - b(q,l), 0\} \tag{16}$$

$$s(i,j,q) = c(i,j) - c(i,q) - c(q,j) + \sum_{l=1}^{n} g(\alpha,q,l) \tag{17}$$

The VAH procedure can be described as follows.

<Vehicle Allocation Heuristic>

1. Initial Tour Constitution

Step 1-1. Find the market $i^* \neq v_1$ which sells more products than any other market, resolve ties by choosing i^* to minimize $\sum_{l=1}^{n} b(i,l)$. Initialize α to contain the depot v_1 and i^* in the initial tour (v_1, i^*, v_1).

Step 1-2. Find the market $j^* \notin \{v_1, i^*\}$ which sells more products that are not sold at i^*, resolve ties by choosing j^* to minimize $\sum_{l=1}^{n} b(j,l)$. Update α to contain j^* in the current tour (v_1, i^*, j^*, v_1).

Step 1-3. Repeat Step 1-2 until the demand for all products is satisfied.

Step 1-4. Apply the 3-opt algorithm to the current tour and get a new current tour α.

Step 1-5. Compute $f(\alpha,l)$ and $g(\alpha,q,l)$ for all l and all $q \notin \alpha$.

Step 1-6. Find the market $q^* \notin \alpha$ and two adjacent markets i^*, $j^* \in \alpha$ such that

$$s(i^*, j^*, q^*) = \max\{s(i,j,q) \mid q \notin \alpha,\ i \text{ and } j \text{ are two adjacent markets in } \alpha\} \tag{18}$$

Step 1-7. If $s(i^*, j^*, q^*) \leq 0$, then apply the 3-opt algorithm to α and stop; otherwise, insert q^* between i^* and j^* and update α, $f(\alpha,l)$ for all l, and $g(\alpha,q,l)$ for all l and $q \notin \alpha$. Go to Step 1-6.

2. Tour Improvement

Step 2-1. For each $i \in \alpha$, compute the increase of purchase cost if market i is to be excluded from the α:

$$g(i) = \sum_{l=1}^{n} [\min\{b(j,l) \mid j \in \alpha, j \neq i\} - \min\{b(j,l) \mid j \in \alpha\}] \tag{19}$$

Step 2-2. Find three adjacent markets i_+^*, i^*, and i_-^* in α such that

$$s(i_+^*, i^*, i_-^*) = \max\{c(i_+, i) + c(i, i_-) - c(i_+, i_-) - g(i) \mid i_+,\ i \text{ and } i_- \text{ are three adjacent markets in } \alpha\} \tag{20}$$

Step 2-3. If $s(i_+^*, i^*, i_-^*) \le 0$, then apply the 3-opt algorithm to α and stop; otherwise, drop market i^* from α by connecting the two markets i_+^* and i_-^* and then go to Step 2-1.

3. Vehicle Allocation

Step 3-1. We are given $p(\alpha, l)$, the purchase priority of product l at any market in α. If a product l can be purchased at only one market, then $p(\alpha, l)$ =1.

Step 3-2. If a product l is the cheapest one in the same product group in α, then $p(\alpha, l)$ =2.

Step 3-3. Give a purchase priority number to all remaining products such that the cheaper ones get the lower priority numbers.

Step 3-4. Define $i(i, l)$ as the inventory of product l at market $i \in \alpha$. Then $r(i)$ could be defined as (21).

$$r(i) = \sum_{l=1}^{n} [i(i, l) \mid i \in \alpha \text{ and } \{p(\alpha, l) = 1 \text{ or } 2\}] \tag{21}$$

Step 3-5. Compute $r(i)$ for all $i \in \alpha$.

Step 3-6. Allocate a vehicle to the market i in α for which the value of $r(i)$ is largest.

Step 3-7. Repeat step 3-6 until all the remaining vehicles are allocated.

Step 3-8. Allocate an unvisited market $q \in \alpha$ next to visited market $i \in \alpha$ which is the nearest one from q.

Step 3-9. Repeat step 3-8 until all the remaining unvisited markets are allocated.

Step 3-10. Complete the each vehicle's tour returning to a depot.

Step 3-11. Purchase a product l for which $p(\alpha, l)$ is 1 or 2 within the vehicle's capacity. If a vehicle's current capacity does not exceed the given capacity then purchase product l, preferentially, that has lower purchase priority number. Repeat the purchasing of products until the demand for all products is satisfied.

Step 3-12. Calculate the total travel cost, purchase cost, and the sum of these.

4. EXPERIMENTAL ANALYSIS

4.1 Testing environment

The conditions of the experiment are as follows. At first, we randomly generated the coordinates of markets, including the depot, according to uniform distribution in a 100×100 2-dimensional plane. We set the number of products n equal to the number of markets m $(m = n)$. In the experiments for small size instances, we set m (or n) $\in \{15, 20, 25\}$ and calculate optimal solutions by using ILOG CPLEX (version 11.1). We also calculate near optimal solutions by VAH algorithm, and compare these with the optimal solutions. In the experiments for large scale instances, we set m (or n) $\in \{50, 75, 100\}$ and calculate near optimal solutions only. This approach is taken because TPP (or MTPP) is NP-hard. Therefore, the calculation of the optimal solution is impossible for the large scale problems. 1/3 of q_{il} is given as 0, and 2/3 as 1. We also randomly generated the b_{il} values in [30, 50] when q_{il} is not 0.

4.2 Experiments for small scale problems

In the experiments for small scale problems, we calculate the objective function values, computational times, and number of markets in optimal tour for some instances. The travel cost between same market is then given as big M and the vehicle capacity is fixed as 10. Table 1 shows the results. As shown in Table 1, the computational time is increasing in accordance with augmentation of the number of markets and products.

Table 1. Experiment results for the small scale problems

Size				Optimal Solution (CPLEX)		
m	n	Total Demand	Number of Vehicles	Number of Markets in Optimal Tour	Objective Function Value	Computing Time (Sec.)
15	15	15	2	4	654	2.75
	20	20	2	5	829	14
	25	25	3	4	1,045	10.76
20	15	15	2	4	618	3.25
	20	20	2	4	782	32
	25	25	3	4	1,003	40.75
25	15	15	2	4	678	172.71
	20	20	2	5	832	235.28
	25	25	3	7	1,047	2,389.12
Mean	-		-	-	-	322.3

For the same problems, near optimal solutions are calculated by using VAH. These experiments are conducted with a PC Pentium4 2 GHz and Visual C++ (version 6.0). Table 2 shows the results. In Table 2, the "Gap" result is defined in equation (22). As shown in Table 2, mean gap value is 4.35%, with the minimum 1.5% and the maximum 6.1%. The mean computing time is 1 second, which is very short time as compared with the result by CPLEX. We can conclude that the performance of VAH is verified because the mean gap value, 4.35%, is acceptable. We also acknowledge that VAH is very effective algorithm from the aspect of the computational time.

Table 2. Gaps between optimal and near optimal solutions

m	15			20			25			Mean
n	15	20	25	15	20	25	15	20	25	
Gap(%)	4.7	5.5	3.3	5.9	5.8	4.5	6.1	1.9	1.5	4.35
Computing Time(Sec.)	1	1	1	1	1	1	1	1	1	1

$$Gap(\%) = \frac{near\ optimal\ solution\ -\ optimal\ solution}{optimal\ solution} \times 100 \quad (22)$$

4.3 Experiments for large scale problems

For the large scale problems, the vehicle capacity is fixed at 30 and other parameters are the same as those in the small scale experiments. The results are as shown in Table 3. Near optimal solutions for some instances are calculated by VAH within just 2 seconds. It's possible to say that VAH is very effective for the large scale problems as well.

Table 3. Experiment results for the large scale problems

Size				Near Optimal Solution (VAH)	
m	n	Total Demand	Number of Vehicles	Number of Markets in Final Tour	Computing Time(Sec.)
50	50	50	2	11	1
	75	75	3	12	1
	100	100	4	14	1
75	50	50	2	15	1
	75	75	3	14	1
	100	100	4	15	1
100	50	50	2	17	2
	75	75	3	16	1
	100	100	4	16	1

5. CONCLUSION AND FUTURE WORK

The purpose of this paper is to present a mathematical formulation of the UMTPP and to suggest a method for finding the solution. Through numerical experiments, we have found out that the suggested VAH algorithm is very effective from the point of view of the computational time. It's also acceptable from a point of view of the gap with optimal solution. Most of related studies for the TPP only consider a single purchaser. Therefore, all of the methods of solving the TPP are not eligible for solving the MTPP. The most important contribution of this paper is to suggest the solution method of solving the MTPP for the first time.

In the future, the development of an algorithm to be eligible for the CMTPP is strongly needed. Also, in this paper, we generated some specific problems. However, if the real data from any enterprise could be used in these kinds of experiments, it would better demonstrate its value.

6. REFERENCES

Angelelli, E., Mansini, R., Vindigni, M. (2009), Exploring greedy criteria for the dynamic traveling purchaser problem, *Central European Journal of Operations Research* 17(2), 141-158.

Boctor, F.F., Laporte, G., Renaud, J. (2003), Heuristics for the traveling purchaser problem, *Computers and Operations Research* 30, 491-504.

Bontoux, B., Feillet, D. (2008), Ant colony optimization for the traveling purchaser problem, *Computers and Operations Research* 35, 628-637.

Burstall, R.M.(1966), A heuristic method for a job sequencing problem, *Operational Research Quarterly* 17, 291-304.

Buzacott, J.A., Dutta, S. K. (1971), Sequencing many jobs on a multipurpose facility, *Naval Research Logistics Quarterly* 18, 75-82.

Choi, M. J., Lee, S. H. (2009), Periodic heterogeneous multiple traveling purchaser problem for refuse logistics optimization, *Journal of the Korean Society of Supply Chain Management* 9(2), 147-154.

Choi, M.J., Lee, S.H. (2010), Heterogeneous multiple traveling purchaser problem with budget constraint, *Journal of the Korean Operations Research and Management Science Society* 35(1), 111-124.

Current, J., Revelle, C., Cohon, J. (1984), The shortest covering path problem : an application of locational constraints to network design, *Journal of Regional Science* 24, 161-183.

Goldbarg, M.C., Bagi, L.B., Goldbarg, E.F.G. (2008), Transgenetic algorithm for the traveling purchaser problem, *European Journal of Operational Research* 199, 36-45.

Golden, B.L., Levy, L., Dahl, R. (1981), Two generations of the traveling salesman problem, *Omega* 9, 439-445.

Jain, R., Rivera, M.C., Moore, J.E., Lake, J.A. (2003), Horizontal gene transfer accelerates genome innovation and evolution, *Molecular Biology and Evolution* 20, 1598-1602.

Laporte, G., Riera-Ledesma, J., Salazar-Gonález, J.J. (2003), A branch and cut algorithm for the undirected traveling purchaser problem, *Operations Research* 51(6), 142-152.

Lin, S. (1965), Computer solutions of the traveling salesman problem. *Bell System Tech. J.* 44, 2245-2269.

Lomnicki, Z.A. (1996), Job scheduling, *Operational Research Quaterly* 17, 314-316.

Mansini, R., Tocchella, B. (2009), The traveling purchaser problem with budget constraint, *Computers and Operations Research* 36, 2263-2274.

Margulis, L. (1992), Symbiosis in cell evolution, Microbial communities in the archean and proterozoic eon., W.H. Freeman.

Miller C.E., Tucker, A.W., Zemlin R.A. (1960), Integer programming formulation of traveling salesman problems, *Journal of Association for Computing Machinery* 7, 326-329.

Ong, H.L. (1982), Approximate algorithm for the traveling purchaser problem, *Operations Research Letters* 1, 201-205.

Pearn, W.L., Chien, R.C. (1998), Improved solutions for the traveling purchaser problem, *Computers and Operations Research* 25, 879-885.

Ramesh, T. (1981), Traveling purchaser problem. *Journal of the Operational Research Society of India* 18, 78-91.

Ravi, R., Salman, S. (1999), Approximation algorithms for the traveling purchaser problem and its variants in network design, *Lecture Notes in Computer Science* 1643, 29-40.

Riera-Ledesma, J., Salazar-Gonález, J. J. (2005a), A heuristic approach for the traveling purchaser problem, *European Journal of Operational Research* 162, 142-152.

Riera-Ledesma, J., Salazar-Gonález, J.J. (2005b), The biobjective traveling purchaser problem, *European Journal of Operational Research* 160, 599-613.

Riera-Ledesma, J., Salazar-Gonález, J.J. (2006), Solving the asymmetric traveling purchaser problem, *Annals of Operations Research* 144, 83-97.

Singh, K.N., van Oudheusden, D. L. (1997), A branch and bound algorithm for the traveling purchaser problem, *European Journal of Operational Research* 97, 571-579.

Teeninga, A., Volgenant, A. (2004), Improved heuristics for the traveling purchaser problem, *Computers and Operations Research* 31, 139-150.

Voß, S. (1996), Dynamic tabu search strategies for the traveling purchaser problem, *Annals of Operations Research* 63, 253-275.

FLEXIBLE PRODUCT ALLOCATION IN DISTRIBUTION PROCESSES IN AN APPAREL SUPPLY CHAIN

Bernd Scholz-Reiter, Michael Teucke*, Anne Schweizer, and Mehmet-Emin Özsahin

BIBA – Bremer Institut für Produktion und Logistik GmbH
at the University of Bremen
Hochschulring 20
D-28359 Bremen, Germany
e-mail: {bsr|tck|vir|oez@biba.uni-bremen.de

Abstract: In apparel supply chains retailers often order large numbers of different product variants with only few pieces, increasing stock levels, while product sales times are often short. This article outlines a concept, how decentralized decision-making based on the concept of autonomous control can be used to achieve more flexible allocation of articles to customer orders. The decision making can use simple heuristic rules. Several rules for articles-to-order allocation are discussed in the context of typical apparel order policies. The rules are applied to a case study in form of an Asian-European apparel supply chain. Additionally, the article examines the requirements of flexible articles-to-order allocations on execution of intra-logistic processes within nodes of the case study's distribution network, such as factories, hubs and warehouses. A technical infrastructure based on the use of RFID technologies is proposed for these network nodes.

1. INTRODUCTION

In apparel supply chains production often takes place in low-wage countries, while demand is located in the industrialized world. Supply chains have to bridge large geographic distances, resulting in long lead times. In contrast, retailers request short lead times and high delivery flexibility, while frequently ordering small numbers of garments across a large variety of product variants. These conflicting objectives often result in high uncertainty and complexity, in particular with seasonal order cycles for fashion products (Brun and Castelli 2008).

One approach to deal with complexity of logistic systems is autonomous control of logistics processes. The concept of autonomous control is characterized by decentralized decision-making by logistic objects (Scholz-Reiter et al. 2009). This article aims to describe, how use of this concept may allow more flexible allocation of articles to customer orders within apparel supply chains. Different heuristic rules for articles-to-order allocation, considering matching of order quantities for product variants, are discussed for typical apparel business/delivery models. The rules are applied to a case study in form of an Asian-European apparel supply chain. The supply chain includes production plants in East Asia and distribution centers in Europe, which deliver garments to retailers in Europe and the Middle East.

Additionally, the paper examines the requirements of flexible articles-to-order allocations for execution of packing and distribution processes within nodes of the case study's distribution network, such as logistic hubs and distribution centers. A technical infrastructure based on the use of smart label technologies is proposed and problems relating to use of RFID systems are discussed.

The structure of the article can be outlined as follows: The second section will describe supply chain and inventory management in the apparel industry, focusing on different order and delivery models as well as common problems in apparel logistics. The third section will introduce the case study of an apparel supplier producing in Asia and distributing the products to customers in Europe. The fourth section will describe autonomous control and how this concept may be applied in order to allocate articles to customer orders more flexibly. The fifth section will examine the requirements of flexible articles-to-order allocations for execution of packing and distribution processes within the nodes of the case study's distribution network. For these network nodes a technical infrastructure based on the use of RFID technologies is proposed and problems relating to use of RFID systems across different continents are discussed. The paper closes with a conclusion and an outlook to future research topics.

2. SUPPLY CHAIN AND INVENTORY MANAGEMENT IN THE APPAREL INDUSTRY

The so called textile process chain includes all process steps of the textile and apparel production and distribution process. The main steps of the textile and garments production process are manufacture of fibers (yarns and threads) by the fiber industry, forming of textile surfaces (fabrics) as well as dressing and colorizing of the surfaces by the textile industry, and manufacturing and finishing of ready-to-wear garments by the apparel industry. The last part of the textile chain is formed by garment retailers, who are responsible for commercial distribution, selling the ready garments to end customers. These steps are usually executed in strictly sequential order, as shown in Figure 1, (Hurcks, 1993). This contribution looks at the part of the textile process chain concerned with apparel production and distribution to retailers.

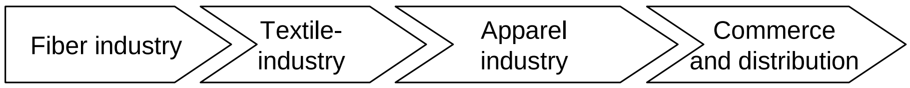


Figure 1. Composition of the textile chain (Hurcks, 1993)

As an important consumer oriented industry, apparel suppliers and retailers have been a driving force to adopt efficient consumer response concepts, like collaborative planning, forecasting and replenishment, or vendor managed inventories (Seifert, 2002, von Heydt, 1999). In reaction to different kinds of products and to the different kinds of retailers and customer requirements, different order and delivery models have been established in apparel supply chains involving retailers and manufacturers or suppliers. The most common order and delivery models are listed in Table 1. Selection of the appropriate model depends on the longevity of the articles' sales periods, which can be grouped into seasonal articles and standard articles sold for many seasons, and according to customer classes. Customer classes can be grouped into small retailers with only one or few sales points, like boutiques or millineries, and large retailers, like clothing chains with many sales points, or mail order wholesalers.

Table 1. Order and delivery models in apparel supply chains (Ahlert and Dieckheuer, 2001, Bruckner and Müller, 2003)

Type of product	Delivery strategy	customer class	
		small retailers	large retailers
Standard products	make-to-stock	• warehouse stock based delivery • never-out-of-stock delivery	• never-out-of-stock delivery
Seasonal articles	make-to-order	• classical seasonal order business	• large retailer orders
Intermediate articles	hybrid forms	• Seasonal Filling Up	• Seasonal Filling Up

As is shown in Table 1, the different order and delivery models can be divided into make-to-stock and make-to-order delivery strategies. Make-to-stock strategies are largely employed for standard articles, which are sold for several seasonal periods without being modified. Warehouse delivery and never-out-of-stock delivery are examples of make-to-stock strategy applications in the apparel industry. Make-to-order strategies are normally employed for more short lived, seasonal products, which are created and sold only for one season and thus follow recent fashion trends. Classical seasonal business and its adapted forms in large retailer business are examples of make-to-order strategy applications in the apparel industry. The most important make-to-stock delivery strategy is never-out-of-stock delivery.

Classical seasonal business is characterized by fixed seasonal cycles with fixed scheduled dates, or phases, for product offers, orders, and delivery, which are repeated twice or four times a year. Assortments of samples, which have been designed by the supplying apparel company for its own brands, are presented to customers by salesmen or at fairs some six to eight months before proposed delivery dates. During a defined order period of roughly three months retailers can order articles in the quantity and variants. Delivery of the products is scheduled by the apparel company at a fixed date or within a fixed time span, starting some three months after the end of the pre-order phase. As production and distribution lead times normally exceed the length of the delivery times expected by customers, production planning and procurement of raw materials have to start before the end of the order periods, using forecasts of the total order sizes based on the orders arrived so far. Products, of which particularly high numbers have been ordered, may be produced in larger quantities and offered again in a second, post-order, phase (Ahlert and Dieckheuer, 2001). As retailers are restricted to ordering a certain quantity of articles from an existing range of products, classical seasonal order is directed largely to small retailers, like millineries or boutiques. For this reason classical seasonal ordering is often characterized by large numbers of orders with relatively small volumes.

For customers, who order large quantities of article items , like clothing chains or mail order wholesalers, the original classical seasonal order business has been adapted to suit their special needs as well as their magnified importance for the supplier's economic success. For this reason, a couple of modifications in dealing with orders of these customers separate them from classical seasonal business. As these customers require specialized products for their own trademarks and, suited to their individual needs product development often is performed as a collaborative process. Ordering and delivery dates are largely dependent on customer wishes as well (Fissahn, 2001).

Most important among make-to-stock delivery strategies is never-out-of-stock delivery, which is a form of vendor managed inventories (Heydt, 1999) employed for non seasonal, long term sold products. Many retailers only hold base stock inventories of these articles, while the suppliers hold inventories to fill them up with short lead times. Retailers review daily (or occasionally weekly) sales volumes in their sales points and reorder these overnight at the suppliers warehouses. Delivery times from the warehouse to the customer are ranging from one to three days. Complete availability (100% service levels) is granted within certain order volume boundaries. Supplier warehouses are refilled periodically from production with a more long termed replenishment process (Ahlert and Dieckheuer, 2001).

Suppliers also occasionally produce articles entirely on their own forecasts, without any pre-orders, and then sell them directly from their warehouses. As long as stocks have not run out, delivery times are comparable to never-out-of-stock delivery. Once warehouse stocks have been used, no reorder possibility for retailers exists (Fissan, 2001).

Several hybrid order and delivery strategies exist, combining different features of the described order and delivery strategies. One example is the so called seasonal filling up products, which combine large retailer orders and warehouse delivery or never-out-of-stock delivery of products during a summer or a winter season. For the next half year, existing retailer stocks are exchanged for the complementary season's products and stored at supplier warehouses, to be delivered again, when the same season arrives again the next year.

Apparel logistics has to cope with a number of specific problems. The most important among these are high product variation, rising stock levels, short product marketing and sales periods, limited demand predictability, and comparatively high product values:

- *High product variation*: Many important retailers have established their own labels and trademarks, for which specific products have to be supplied. These products have to be adapted to the individual requirements of the customer and cannot be supplied to other customers. Additionally, each garment type can be delivered in many different variants, which differ in their colors, sizes and prints. This results in large numbers of brands, fabric qualities, cuts, colors and sizes, which can be combined in various ways. Customers often order small numbers of garments across a large variety of product variants.
- *High and still rising stock levels*: For never-out-of-stock delivery in particular, suppliers often have to keep large inventories in order to be able to react to demand fluctuation faster than with the mentioned, long delivery times. Stock levels in general are still rising due to growing product diversification and variation resulting in rising numbers of products and product variants, to be kept in stock (Scholz-Reiter et al., 2009).
- *Large share of products with short marketing and sales periods*: Many product lines are marked by regular, seasonal exchanges or updates of the product assortment. Once their season has ended, many of these seasonal articles can only be sold with large price reductions or not anymore at all.
- *Volatile demand*: Demand fluctuations are particularly difficult to predict for many short lived, in particular seasonal, products in apparel supply chains due to the effects of fashion tastes. Delivery dates have to be synchronized with anticipated changes between summer and winter collections in stores. The exact dates of the changes however depend on seasonal weather changes, which are difficult to predict during production planning. Retailers, who have to react flexibly to modifications in customer requirements, require from their suppliers the same flexible reaction to demand fluctuation.
- *High product value*: Containers and storage areas filled with garment pieces, represent a comparatively high value for the suppliers.

Retailers additionally have to cope with serious theft rates and relatively high product replacement for end customers unsatisfied with their original buying. These conflicting situations often result in high uncertainty and complexity, in particular for fashion products with short marketing times (Brun and Castelli, 2008, Tellkamp and Quiede, 2005).

As make-to-order products are more challenging in apparel logistics, distribution in classical seasonal business and large retailer order business will be studied in more detail based on a case study of a German apparel supplier.

3. DESCRIPTION OF THE CASE STUDY

The work domain of the supplier is production and distribution of casual and leisure wear. The product assortment includes own brands as well as trademarks of large retailers. For main brands, up to 90 different product types are developed each season. Each garment type can be delivered in different variants, which differ in their colors and sizes; there are e.g. 64 standardized sizes for different countries. Due to the potential combinations of article types, fabric qualities, cuts, colors and sizes the supplier offers several tens of thousands of different article variants.

Altogether, the supplier delivers more than 5 million garment pieces a year of these variants to roughly 6,000 sales points. One half of the sales points belong to large clothing or retail chains, while the other half are small retailers (millineries and boutiques). Most of the customers are situated in central and Western Europe. There is a small, but growing, share of customers situated in Eastern and Northeastern Europe, the Middle East and East Asia.

The supply chain of the case study includes raw material suppliers and a garment production plant situated in southern China, a procurement agency situated at Hong Kong, and distribution centers situated in Europe (Germany). This supply chain is similar to the generic apparel reference supply chain described by Bruckner and Müller (2003).

Production volumes are based on customer orders and own forecasts, as explained in section 2. The orders are sequenced according to required delivery dates and availability of raw materials. Procurement lead times for raw materials are roughly 40 days. Production order throughput times are roughly 10 days from the arrival of the required raw materials. The ready-made garments are transported in standard containers, packed into cartons. A container load consists of 200 to 500 packages of up to 20 pieces each. Packing of packages and containers is product type and variant specific, however not order specific, as order specific picking is restricted to the warehouse. During packing, articles, which have completed production, are counted manually and documented in packing lists, which serve as arrival forecasts for distribution centers.

At the moment product distribution is executed by routing all garments through the distribution center. The distance between Asia and Europe is normally bridged by ship, but in urgent cases air transport may be used. Transports are executed completely by external service providers and are by default routed through logistical hubs, where transport routes and means can be switched. Transport times by container vessel are 30 to 40 days, depending on the route, which is reasonably close to values found in the literature (Pfohl *et. al.*, 2007). Air transport times are only two or three days, but increases transport costs of the same volume and over the same distance by a factor of three to ten compared to transport by container vessel. After arrival at the distribution centre incoming containers are unloaded and packages are put on palettes and stored at free space storing areas until delivery. The packages are counted and compared to dispatch notes, but the articles are only randomly controlled by opening a small sample of the packages. For this reason packing or counting errors at production are often detected only during later picking or even later, by the customer. If stocks of arriving articles in the distribution centre have already run out or have become insufficient, arriving packages will not be put into store, but will be directed to shipping areas of the distribution center, where picking teams can take from them the ordered pieces.

Current process execution with its article being routed through the warehouse results in a number of problems, which are related to the general challenges described in section 2. Arrivals of garments for a product season or a large retailer order cause short, but large peaks in cumulative warehouse stocks and demand large storage areas. After a short time most of these articles have left the warehouse again to be delivered to the customers. However, as the stock based delivery forms are continued independent of these stock level peaks, they need separate storage areas with storage racks for easy picking. For that reason, the areas needed to store seasonal and large customer orders cannot be used for other purposes. This situation results in poor utilization of a significant part of the warehouse storage space.

Picking of orders at the warehouse requires large manual efforts for opening of packages, withdrawal of pieces of the appropriate product variants in the required quantities, and repacking them. This situation results in sharp work load peaks. As timely throughput of large orders is impossible with the warehouse work force, additional workers have to be hired, who are not used to the warehouse processes. This results in poorer efficiency of work processes and higher risk of errors. Routing of garments destined to customers located e.g. in North Eastern Europe or the Middle East through the warehouse results in additional transports and warehouse times and thus in longer product delivery times. Routing through different countries increases efforts for customs handling and product traceability. Difficulties to export garments produced in China to Russia via the European Union may be mentioned as an example.

For these reasons the supplier plans to restructure distribution to direct delivery of the products from factories to part of the customers. Instead of routing all articles through the warehouse, part of the articles should be delivered to customers directly from the factories, while only part of the articles to be delivered is routed through the warehouse, as illustrated in Figure 2. This should result in a more flexible product routing and delivery processes.

Direct delivery however has to cope with a number of challenges due to short term modifications of customer orders. With warehouse based delivery, these are not as serious, as routing of the articles through the warehouse allows reacting to these circumstances, while the articles have arrived there.

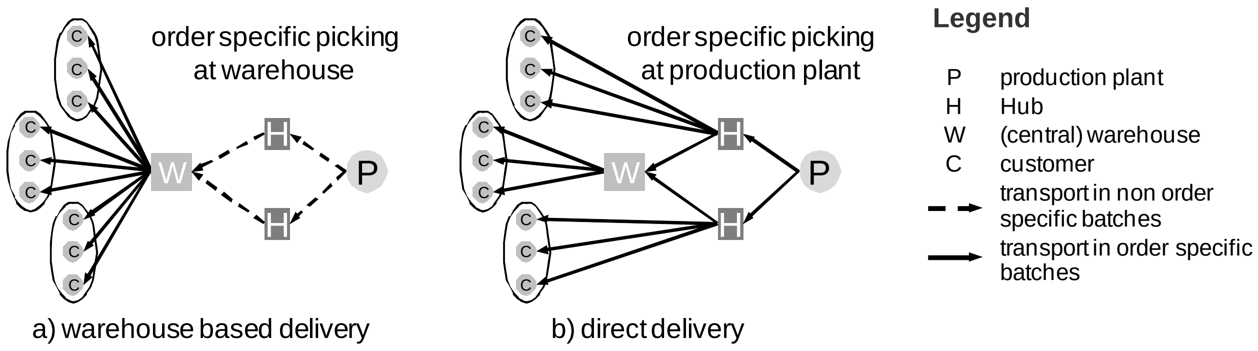


Figure 2. Customer specific article bundling for warehouse based delivery and direct delivery

Customers in classical seasonal business are mainly small millineries and boutiques, which face a high risk of insolvency. Up to some 5% to 10% of the orders have to be kept back, as the ordering customer has become insolvent or is in immediate danger of becoming so. The articles serving such customer orders have to be redirected either to other customers, whose orders are still open, or to the warehouse, where they may be marketed.

Most large retailers however are clothing chains with many, geographically distributed outlets and sales points. Allocation of the cumulated order volumes to the different sales points is often delayed until after the ready-made garments have left factories and are in transport to the distribution center. Allocation may change due to opening of new or takeover of existing outlets, thus increasing overall demand volumes, which require reallocation of articles from existing outlets. Additionally, actually produced article quantities often differ to some degree from ordered quantities, or part of the articles may be damaged or lost during transport. In case of larger differences, more complete reallocations have to be effected between different outlets.

4. AUTONOMOUS CONTROL AND ARTICLE-TO-ORDER ALLOCATIONS

Autonomous control is characterized by processes of decentralized decision making of interacting elements in nondeterministic systems, which possess the capability to render decisions independently. The collaborative research center (CRC) 637, which studies autonomous control at the University of Bremen, postulates that appropriate interaction of the system's elements will improve the performance of a complex logistic system beyond the element's isolated properties; the system will show positive emergence. In particular, application of autonomous control aims to improve a logistic system's response to dynamic instability, caused by sudden changes in the environment, e.g. the market situation, or within the system, e.g. by breakdown of a machine or transport means. This can be summarized by the term increased robustness. (Windt and Hülsmann, 2008).

The elements of such a system are characterized by their ability to process information, to render and to execute decisions on their own. These autonomous, intelligent objects have to be able to collect and process information on their environments and to identify and evaluate alternative process executions, like e.g. alternative transport routes within a logistic network, according to their individual evaluation system. (Böse and Windt, 2007).

Autonomous control is applied to the case study as a simple model, with packaged bundles of article pieces as one kind, and customer orders as a second kind of autonomous logistic objects. Packages will allocate themselves to customer orders based on local interaction between these objects. This allocation then fuels further decisions during the article distribution process, like e.g. selection of the next target nodes for transports of garments within the network, and their transport means, or storage at the factory. As the objects are geographically distributed, the nodes of the supply network will serve as intermediary objects, which collect and disseminate information about the articles and orders. In this way, the allocation can be made for any garment package within the supply network, not only at the final distribution centre. The different types of objects, their roles, objects and tasks are listed in Table 2.

Decisions by the logistic objects have to be based on appropriate decision methods. One potentially viable approach is implementation of rule based decisions (Böse, *et al.* 2005). Such a rule based approach will be followed here.

In a first step, the decision alternatives for each package have to be determined. Potential alternative orders for the allocation of a package can be included into an individual target order set of that package. These alternatives have to satisfy conditions concerning product type and variant (color and size) and delivery dates, compared to current locations and minimum transport times. This order set is gradually reduced during the routing of the articles through the supply network starting with production at the factory, as restrictions become more distinctive and alternatives not meeting these tougher restrictions have to be dropped. In a second step, packages have to select the most appropriate customer order by a package rules for re-allocation of units to orders.

- During production of the garments in the factories the set includes all customer orders, which include articles of the same product type and product variants as those in production. Regional distribution of customers is not important at this stage. This set is uniformly assigned to all garments of the production order. To provide an example, all garments belonging to a production order, which is based on the customer order of a large retailer, will have all sales points of that retailer assigned to them uniformly. No actual allocation of individual garments to customer orders is necessary.
- During packing and dispatch of the ready-made garments packages are allocated to an individual customer order, and filled with garments accordingly. The allocated order is provided the package number and can adjust the number of its "open" articles accordingly. The package's target order set still includes all customers, who have ordered products of the respective type and variants.
- During transport to a hub and arrival at the hub, due to some dynamic influences as described in section 3, a package may be allocated to another customer order or to no order at all and instead redirected to the warehouse. The package's target order set includes all customers, who have ordered products of the type and variants and whose products are routed through the respective hub.
- Dispatch of the garments at a hub and transport from the hub to a customer are the last steps, during which the only one finally selected customer order is allocated to each package.

Table 2. Types of autonomous objects in the model and their main properties

objects role		**objectives**	**tasks**
customer order	representation of a customer order	high adherence to delivery dates	• track keeping of the number of "open" garment pieces of each individual article type, for which no articles are yet allocated; • computation of the order priority, based on the importance of the customer; • track keeping of order status;
package	representation of a package with a distinct number of article pieces of certain product variants	short delivery times	• request of orders of articles of their product type and variant (target order set) from hub object; • interaction with orders belonging to target order set request of article allocations; • selection of a customer order to allocate; • selection of the next node in the supply network, and the respective transport means;
supply network hub	representation of hubs in the supply network (factories, logistics hub), information broker	few reallocations, short storage time	• collection and updating of all orders of those customers (sales points), which are served through the respective hub; • identification of packages and their content; • dissemination to package objects of information on orders with articles of their product type and variants; • computation of transport times to other nodes and dissemination of information to packages; • transfer of the resulting packing lists (transport forecasts) to neighboring hubs;

Figure 3 illustrates for a very simple example the mechanism how a package receives information of its target orders from a node object, then interacts with these orders and allocates itself to one of them, based on application of the slack time rule.

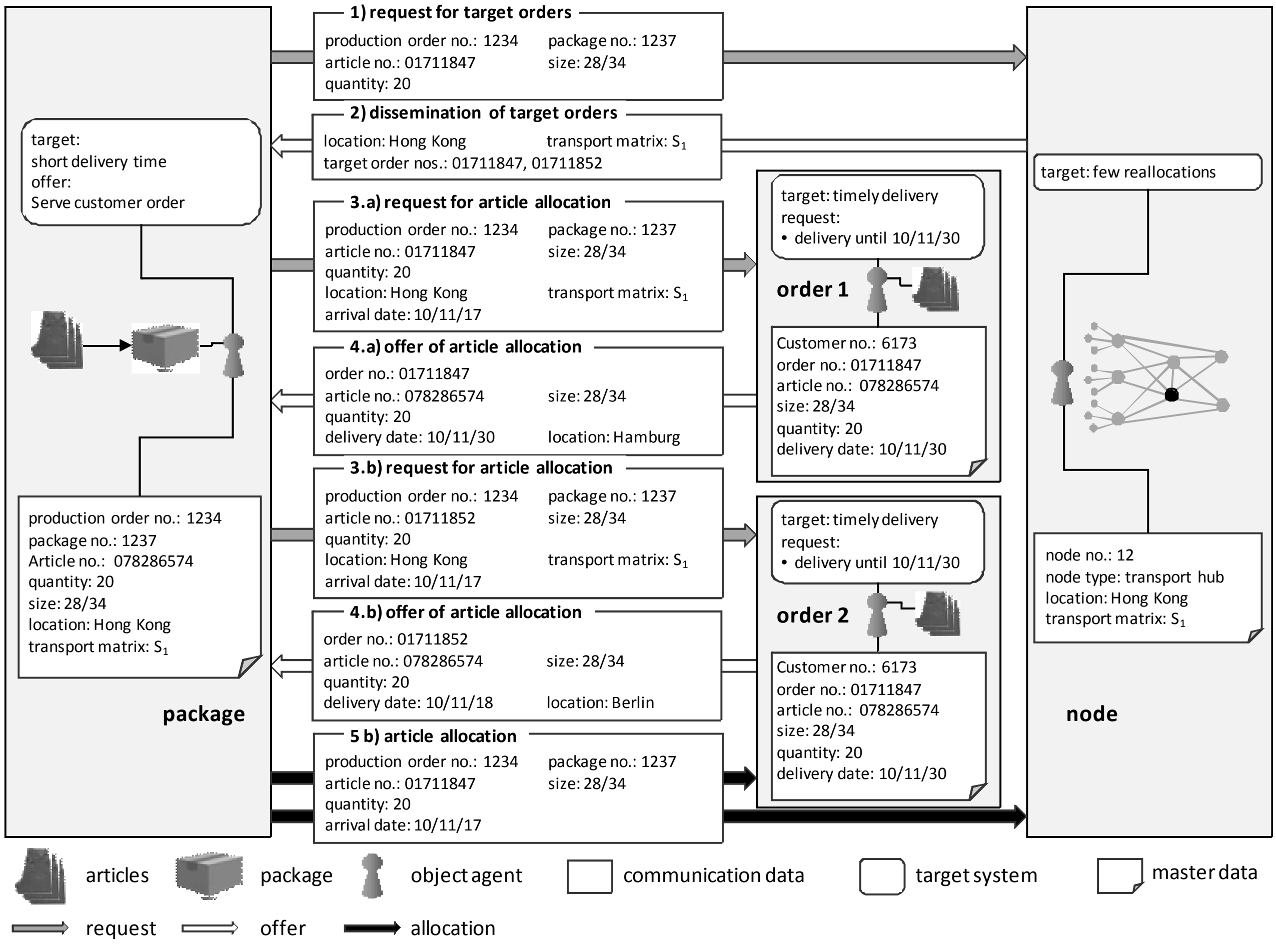


Figure 3. Procedure of package to order allocation based on slack time rule

As mentioned, the actual selection of an order is based on a ranking of all orders of the unit's individual target order set. Based on the requirements of the studied apparel supplier, the following heuristic rules are used to prioritize the customer orders for selection by the packages:

The *shortest servable delivery time* rule is an adaptation of the slack time dispatch rule, which is used for sequencing of production orders in production control (Kistner and Steven, 2001). Packages are allocated to the order with the shortest slack time, which measures the difference between contracted delivery date and anticipated transport times to customer. Slack times until delivery are to be compared to transport times of the articles to the customer. Application of this rule, which is less myopic than alternative dispatch rules, decreases total tardiness for customer orders.

If *customer based order priority* is applied, garments are allocated with priority to those orders of those customers, who are of highest importance for the supplier. Packages with appropriate article variants will be allocated to the highest prioritized order until it is completely served with articles. After that the next highest prioritized order will be served. Customers may be ranked in importance or priority by the supplier according to current or historical overall order volumes, or similar criteria. Application of this rule decreases tardiness for orders by large customers, but may increase tardiness for orders by less important customers. Accordingly, this rule is mainly applied in those cases, when the shortest servable delivery time rule is indifferent to priorities of two different orders.

Generally, differences between ordered and allocated articles have to be minimal over the ordered product variants. The quantities of each product variant is identified in each package and compared to the ordered quantities. Packages are allocated to those orders, where the difference between ordered quantities and physical quantities is minimal over all product variants. Several difference measures can be used, like modulus difference, or quadratic difference, which are summed up for each product variant to arrive at an overall difference.

5. EXECUTION OF PACKING AND DISTRIBUTION PROCESSES

Direct delivery requires individual allotment of ready-made garments to customer orders already at the factories and the hubs. For this reason, articles have to be individually identified at the factories and then tracked during their transport through the supply chain.

RFID technology (Shepard, 2004, Finkenzeller, 2004) offers potential to store individual product related data and additional logistic process data both at item and at unit level, allowing more easily associating correct data to products. This allows better synchronization of physical material flows and associated data flows over supply chains (Gillert and Hansen, 2007, Schuster *et al.*, 2007). Application of transponders at item level may be permanent, either by direct integration of transponders into textile surfaces (weaving, sewing or printing), or by application as distinct, permanently applied, but principally removable labels, which are sewed into textiles. For item level applications, a variety of special transponders have been developed (Kallmayer *et al.,* 2003). However, permanent application is still problematic, as life expectancy of transponders added before finishing may be strongly reduced by washing and ironing processes. Permanent labels may reduce wearing comfort of the garments, and acceptance by end customers is not guaranteed.

Alternatively, application may be temporary in form of easily removable labels or stickers. For garments that are stored and transported on hangers, many solutions for transponder integration into hangers have been developed. These are mostly handled sequentially as single pieces in indoor logistics using automated transport systems like conveyors. As this eases automatic identification, barcode and RFID technology for identification are already widely used, often via permanent or removable transponders on the individual garment packaging foils (Schmidt and Mannel, 2002). Storage and transport of garments without hangers, but folded into packages, reduces transponder application on item level. However, it allows application of cheap standard transponders on unit packages like cartons, which reduces the problems relating to transponder technology and application on or in garments, as well as the number of transponders needed to identify the garments.

For application at the case study, combined use of transponders on item and unit levels has been proposed and tested. This allows individual handling of the garment pieces, when needed, as well as limitation to package identification, as long as sufficient. This way, RFID is used to provide the extra-information to allow decentralized decision-making within the supply network's nodes.

Part a) of Figure 4 shows the transponder labeling and identification process flow at the factory. During the finishing process the ready-made garments are equipped with easily removable transponder labels. The labels can be added as integrated components with the routinely attached size books and price stickers. Based on the Electronic Product Code (EPC) standard, which defines the form and the content of a RFID code, each piece's information, including article number, size and color information, and an additional serial number of the individual piece is stored into the transponder label. During packing of garment pieces into packages, the pieces can be identified and counted, using either mobile RFID readers or RFID tunnel readers. It is checked that all packages contain the defined number of garments. As packages are equipped in the same way with transponder labels, they can be identified with their included articles as well. The identity of the packages can be allocated to the orders in the order database.

After packing the packages are stuffed into a container, which is used for transport between the production plant and the distribution centers or customers. During container stuffing, the packages are identified and counted using either mobile or gated RFID readers. A packing list of all the packages and articles in the container can be automatically created and sent to the target hub. The packing list can also contain information for customs service procedures.

The intra-logistic processes in a logistic hub are shown in part b) of Figure 4. The processes include warehouse entry, redistribution of packages (as far as necessary), and dispatch. During warehouse entry, an RFID-reader reads the information which is stored on the RFID label belonging to each package. The gathered information is stored in the order database. The actual stock level data of the hub is updated; the order database is compared to the packages receiving simultaneously. If the order has been modified or cancelled during transport, the packages destination will be rearranged to other customers, as described. A similar order can be brought forward for instance. The decisions of rearrangement rest upon information in the order database and application of the allocation rules presented in section 4. If the order is still valid, no rearrangement is necessary. RFID-readers at the warehouse read the information of every leaving package while it is put into a container or trailer.

Part c) of Figure 4 sketches the architecture of a demonstration tool, which realizes the article to order allocations based on the principles outlined in this article. The tool offers interfaces for order entry, RFID based identification of single garments and packages, allocation of packages to orders, and creation of packing lists. Instances of the application can be installed in several network nodes. Next to the main components the tool includes several interfaces to RFID reader infrastructure.

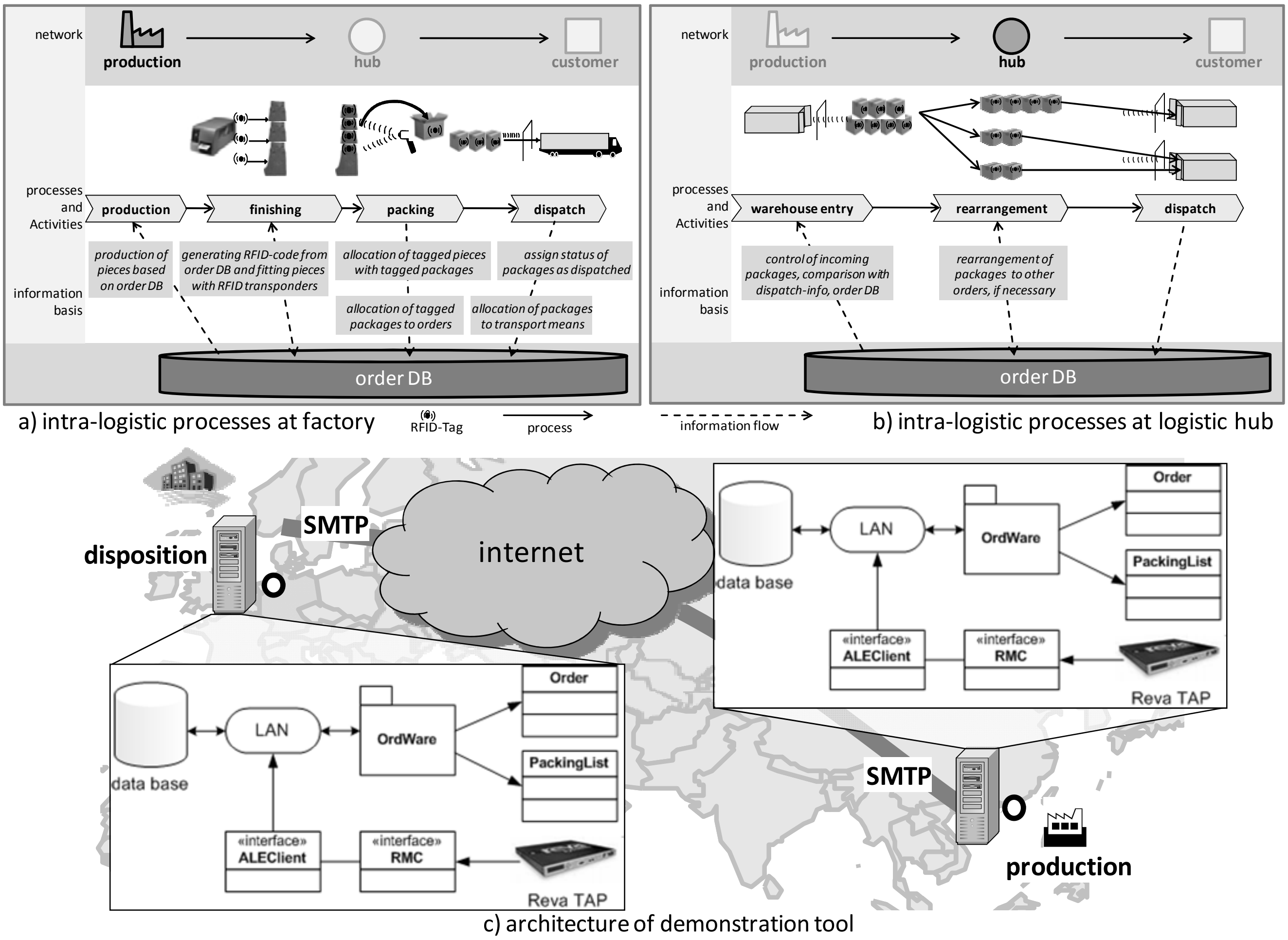


Figure 4. Intra-logistic processes at factory (a) and hub (b); architecture of demonstration tool (c)

6. CONCLUSION AND OUTLOOK

Based on the results of a case study, this paper outlined a concept, how decentralized decision-making based on autonomous control of logistic processes can be used to achieve more flexible allocation of articles to customer orders in apparel supply chains, which have to cope with large numbers of different product variants. Autonomous control can be applied by using packaged bundles of article pieces and customer orders as autonomous logistic objects. Packages will allocate themselves to customer orders based on local interaction between these objects. This allocation then fuels further decisions during the article distribution process, like e.g. selection of the next target nodes for transports of garments within the network or storage at the factory. As the objects are geographically distributed, the nodes of the supply network are used as intermediary objects. In this way, the allocation can be made for any garment packages within the supply network, not only at the final distribution centre. Allocation decisions can be rule based, using combinations of priority rules to rank customer orders, and differences across different product variants. Although the paper deals with a specific case study, results can be adapted to other supply chains adopting similar business models.

Further research is necessary to refine and expand the model of the autonomous logistic objects within transport scenarios and to assign the objects fitting local objectives. The validity of the concept has to be checked by simulations and additional test trials.

ACKNOWLEDGEMENTS

The authors are grateful to the generous support from the German Research Council (DFG) as part of the Collaborative Research Centre 637 “Autonomous Cooperation Logistic Processes – A Paradigm Shift and its Limitations” (CRC637).

7. REFERENCES

Ahlert, D. and Dieckheuer, G. (2001). *Marktorientierte Beschaffung in der Bekleidungsindustrie*. FATM, Münster, Germany.

Böse, F., Piotrowski, J., Windt, K. (2005). Selbststeuerung in der Automobil-Logistik. *IndustrieManagement 20, 4:* 37-40.

Böse, F. and Windt, K. (2007). Catalogue of Criteria for Autonomous Control in Logistics. In: Hülsmann, M. and Windt, K. *Understanding Autonomous Cooperation and Control in Logistics – The Impact on Management, Information and Communication and Material Flow*. Springer, Berlin, Germany, 57-72.

Bruckner, A. and Müller, S. (2003). *Supply Chain Management in der Bekleidungsindustrie*. Forschungsstelle der Bekleidungsindustrie, Cologne, Germany.

Brun, A. and Castelli, C. (2008). Supply chain strategy in the fashion industry: Developing a portfolio model depending on product, retail channel and brand, *Int. J. Production Economics, 116:* 169–181.

Finkenzeller, K. (2004). R FID *Handbook – Fundamentals and Applications in Contactless Smart Cards and Identification*. Carl Hanser Verlag, Munich, Germany.

Fissahn, J. (2001). *Marktorientierte Beschaffung in der Bekleidungsindustrie*. Thesis (PhD). Münster University, Germany.

Gillert, F. and Hansen, W.-R. (2007). *RFID für die Optimierung von Geschäftsprozessen*. Hanser, Munich, Germany.

Heydt, A. von der (1999). Efficient Consumer Response - so einfach und doch so schwer. In: Heydt, A. von der (Ed) *Efficient Consumer Response – Konzepte – Erfahrungen und Herausforderungen*. Franz Vahlen GmbH Munich, Germany, 2–23.

Hurcks, K. (1993): *Internationale Beschaffungsstrategien in der Textil- und Bekleidungsindustrie*. Thesis (PhD). Münster University, Germany.

Kallmayer, C., Pisarek, R., Neudeck, A., Cichos, S., Gimpel, S., Aschenbrenner, R. and Reichlt, H. (2003): New assembly technologies for textile transponder systems. *Proceedings of the IEEE Components, Packaging, and Manufacturing Technology Society: 53rd Electronic Components & Technology Conference*. IEEE Service Center, Piscataway, NJ, USA, 1123-1126.

Kistner, K.-P. and Steven, M. (2001). *Produktionsplanung*. Physica-Verlag, Heidelberg, Germany.

Pfohl, H.-C., Gomm, M. and Shen, X. (2007). China: Textil- und Bekleidungs-Supply Chain zwischen Deutschland und China. In: Wolf-Kluthausen, H. (Ed*). Jahrbuch der Logistik 2007*. free beratung, Korschenbroich, Germany, 258-264.

Schmidt, J. and Mannel, A. (2002). *Einsatzpotenziale der Transpondertechnologie in der Bekleidungsindustrie*. Forschungsgemeinschaft Bekleidungsindustrie, Cologne, Germany.

Scholz-Reiter, B., Teucke, M., Özsahin, M.-E. and Sowade, St. (2009). Smart Label-supported Autonomous Supply Chain Control in the Apparel Industry. *Proceedings of the 5th international Congress on Logistics and SCM Systems,* Program Committee the 5th International Congress on Logistics and SCM Systems, 44-52.

Schuster, E., Allen, S., Brock, D. (2007). *Global RFID: the value of the EPCglobal network for supply chain management*. Springer, Berlin, Germany.

Seifert, D. (2002). *Collaboration Planning Forecasting and Replenishment*. Bonn: Galileo Press.

Shepard, S. (2005). *RFID - Radio Frequency Identification*. The McGraw-Hill Companies, New York.

Tellkamp, C. and Quiede, U. (2005). Einsatz von RFID in der Bekleidungsindustrie – Ergebnisse eines Pilotprojekts von Kaufhof und Gerry Weber. In: Fleisch, E., Mattern, F. (Eds): *Das Internet der Dinge*. Springer, Berlin, Germany, 143-160.

Windt, K. and Hülsmann, M. (2008). *Understanding Autonomous Cooperation and Control in Logistics – The Impact on Management, Information and Communication and Material Flow*. Springer, Berlin, Germany.

INTRODUCING THE CONCEPT OF AUTONOMOUS AGENTS AS A TOOL TO ASSIST THE ANALYSIS OF THE COLLABORATION BETWEEN MANUFACTURED EXPORTERS AND MARITIME SHIPPERS

Vanina Macowski Durski Silva[1*], Antônio G. Novaes[1], Bernd Scholz-Reiter[2], and Jakub Piotrowski[2]

[1]Department of Production Engineering
Federal University of Santa Catarina, Florianópolis, SC, 88040-970, Brazil
e-mail: vaninadurski@gmail.com
e-mail: novaes@deps.ufsc.br

[2]Bremen Institut for Production und Logistic
BIBA- University of Bremen, Hochschulring 20, 28359, Bremen, Germany
e-mail: bsr@biba.uni-bremen.de
e-mail: pio@biba.uni-bremen.de

Abstract: Along export/import chains transportation has an important cost impacting directly on the efficiency of the whole chain. Experiments show satisfactory results in terms of reduced delivery time, increased productivity of transportation resources as well as economies of scale by the implementation of the Collaborative Transportation Management (CTM). It aims to create business alliances oriented to share resources and information. In this context this study proposes a concept of a tool for helping companies on their decision making. It identifies the major parameters that could support the maritime logistics of manufactured goods and, based in the autonomous agent theory applied to logistic problems, the problem could be modeled. As a result, is intended to discuss the new concept of CTM and autonomous agents applied to maritime logistic and manufactured export companies in the way to contribute with companies in order to reduce costs in their chains.

1. INTRODUCTION

In Brazil, the foreign trade has been not used as pro-active factor in development strategy because, historically, the negotiations between the different participants of the export chain have presented conflicts. It is observed that each link intend to minimize its individual costs, which normally does not converge to the global optimum of the supply chain. Therefore, companies are being obliged to re-analyze its procedures, to use reengineering techniques and redefine the relations and models of its supply chains in order to reduce costs, increase efficiencies and gain competitive advantage.

To reduce such problems it has recently emerged the concept of CTM, in the new concept of collaborative logistics. It has been spread out from the year 2000 through Collaborative Planning, Forecasting and Replenishment (CPFR) approach, and CTM has been defined by experts as a helpful tool to provide reductions in the costs of transactions and risks, enhance the performance of service and capacity, as well as the achievement of a more dynamic supply chain (Silva *et al.*, 2009).

As the exporter Brazilian companies are looking for higher competitiveness, they shall not act in an individual manner and start acting in a collaborative manner. Therefore, it is required a detailed sharing of data and information by the agents of the logistics chain, in order to compose a solid partnership. It is understood as agent each integrant of this chain, as in the maritime logistics chain: the producer company, road transportation, shipowners and maritime shippers.

After bibliographic studies and contacting entrepreneurs of this area, it is verified that there is restrict scientific work exploring this subject comprising manufactory industries, freight contractors and maritime shippers, in order to contribute with exportation. Therefore, this study, which is part of a Ph.D. thesis in development, intends to summarily present an overview of the Brazilian exportation and its operation of manufactured exportation chain using maritime shippers, the definition of the new CTM logistic term, its possible implementation in a maritime logistics chain, as well as to begin the application of the concept of autonomous agents to model and analyze this transportation problem.

This is a bibliographic work and, regarding its nature, it can be considered qualitative, aiming at generating knowledge for further practical application in logistic problems. The article is composed by 6 items, including this introduction, distributed as follows: the second item presents a short overview of the Brazilian exportation, the third item contextualizes the CTM approach and the fourth item presents the use of the concept of Autonomous Agents applied to logistic problems. The fifth item presents the prototype model of resolution of the maritime transport logistic and the sixth item presents the final considerations, as well as suggestions for the continuity of this study.

2. EXPORTATION OF MANUFACTURED PRODUCTS

The Economic Commission for Latin America and the Caribbean (CEPAL) classifies the exportable items in manufactured products, basic products and semi-elaborated in function, generally, of the used resources, work intensity, scale intensity and science applied to manufacture. Manufactured product is the one resulting from a standard serial production process, by machines, tools and labor, currently also treated as industrialized product.

In the next topic it shall be presented the scenario of the Brazilian manufacturing sector, including applied volumes and figures, as well as other relevant factors such as barriers to the expansion of its exportation and the presentation of the mechanism when performing the exportation of products using maritime transportation (which is the focus of this work).

2.1 Global demand overview

In the year 2008, the global trades increased about 15%, where 60% of this percentage was concentrated in developed countries. Latin America has participated with 5.4% and Mercosul with only 1.7% of exportations. Regarding the sector distribution of the global exportations of the year of 2007, the latest statistic information available, prevail the products of higher aggregated value - machines, telecommunication equipment, transport, etc. – with a participation of 35.5% of the value. The Brazilian sales, however, represent only 0.7%.

Besides the relevant contribution of manufactured products for the Brazilian economy, there are several problems providing the representativeness of only 0.7% of the total amount of manufactured exported worldwide. One example to mention is the case of barriers imposed to the exportation chain, which may occur externally and internally. Regarding external barriers, it shall be highlighted: requirements of technical standardization and of technology nature, restrictions of environment nature and fusions of companies or intercompany agreements, etc. Regarding internal barriers, we can say those are resultant from a deformed vision regarding the importance of foreign trade of goods and services for the sustaining of economic growth, as a factor for the increase of job offers, induction of investments and motivation for the incorporation of new technologies.

2.2 Mechanism for the performance of maritime transportation in case of exportation

When deciding to export, the company shall observe the stages of the process, aiming at knowing its client market, its demands, habits and characteristics. It shall also define other points: the good transport mode, proper packing of the product in order to maintain its integrity, freight form to be adopted in the negotiation, as well as the company to perform the transportation, besides considering or not the support of an intermediate agent (freight forwarder) or NVOCC (on Vessel Operator Common Carrier) in the negotiation.

Considering that the most part of the exportations is performed through maritime transportation and that this work intends to study this topic, it shall only discuss maritime exportation. Figure 1 shows shortly the stages of the exportation mechanism.

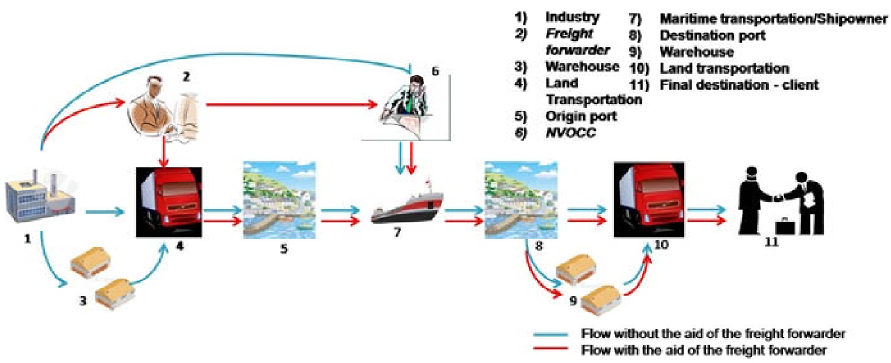


Figure 1. Stages of exportation mechanism
Source: author

In maritime transportation the most important intervenient agents are the industries, land carriers shipowner, maritime agency, NVOCC, load freight contractor, multimodal transport operator and cargo broker. Upon the performance of a sale or purchase it shall be established a delivery point of the goods, where the liabilities shall be shared between the vendor and the purchaser. These liabilities comprise costs and risks on the transaction, and the exporter and importer shall undertake them until the delivery point and from there on.

2.2.1 Transport negotiation

In case the alternative is sending goods as free general load, or even in pallets, the exporter or its agent (freight forwarder) shall be able to perform booking of a conventional ship of regular line (trough a NVOCC or not); or even freight a ship in case of a larger batch.

The next point is to decide through which port the goods will be shipped, as well as the terminal and shipowner to be used. Next, it shall be define the availability mean of the goods in the chosen terminal based on the cost of transportation, traffic time and deadline for the delivery of the goods. One possibility to be considered is the convenience to have a safety warehouse in another city, which means, to maintain a stock in some point to solve quick delivery problems or small quantities which cannot be always quickly produced or sent.

As a consequence of the planned logistics it is possible to reduce the costs of transportation, storage, loss of time on traveling and, finally, problems in the delivery and compliance with sales contract. Therefore, in the decision regarding the port to be chosen by the exporter it shall be also considered the quantity of goods to be produced and shipped.

It is important to mention the real situation where frequently the shipowners decide in a *Freight Conference* the prices to be practiced in the maritime market. In these situations, industries acting alone do not have enough power to better negotiate with them. This is the place where the collaboration can be applied in order to create groups of industries with the same goal, negotiating with the shipowners in order to stop the oligopoly created by them.

Regarding the negotiation rules it is important to mention that each country has its own set of rules for this kind of negotiation (Keedi, 2007), but to avoid problems when two or more countries are negotiating it was created a set of international uniform terms named International Commercial Terms (INCOTERMS).

2.3 How to operate in exportation? Individually or in collaboration?

After defining the general maritime exportation mechanism, it is observed that it is very comprehensive, comprising several variables and requiring updates regarding volumes, capacity of the ships, prices, maritime fees, extras fees, law, among others. One point to be discussed is how the exporter companies involved in this mechanism shall act: individually or in collaboration? If isolated, each exporter company shall analyze the manner to perform the transport of its goods, analyzing the existent navigation companies, fees, acceptable volumes, transport mode of the goods and the proper port, as well as preparing the required documentation.

One possibility to begin collaboration is searching for help from a professional prepared to perform such activities, the freight forwarder. This agent can, if possible, join the loads of other clients in order to achieve scale savings. It can be also performed directly through partnerships among exporter companies, with no contracting of this third party.

Such alliances are under constant transformation, altering its members to adjust to market instabilities or to expand services; those can look like instable associations in a first moment, but are actually flexible and eliminate more fragile competitors, expand the service coverage, enhance efficiency, punctuality and speed. Through this collaboration in the transport it is observed the formation of global logistics networks articulating productive areas, not only expanding service coverage but also frequency, efficiency and circulation rate of the goods (Souza, 2009).

The following item shall approach CTM, in order to explore this new logistics concept and trying to apply it in logistics chain of exportation of manufactured products through maritime shippers.

3. CTM CONTEXTUALIZATION

Considering that the Brazilian GDP was correspondent to R$ 3.14 trillion in 2009 and that the percentage applied to transportation was of 5.4%, we can say R$ 169.72 billion were invested in this segment (Mdic Statistics, 2010). This percentage has decreased but it doesn't represent Brazil is in good shape if compared to developed countries. It shows that the transportation is a significant item in final accountability of resources spent on Brazilian products, either for the domestic market as for products to be exported, composing the "Brazil Cost" (Tacla, 2003).

One manner to reduce the existing problems in logistics, more specifically in transportation, is trough collaboration between the several agents of a logistics chain, adding value to the whole process. It is also valid to mention that, for more benefits, the CTM shall act in the interaction of several logistics chains, forming the concept of logistics network. For Tacla (2003) the collaboration is performed through sharing resources, mainly in the use of the same transportation equipment, contributing for the reduction of costs by enhancing the productivity of the transportation equipment. Therefore, the beginning of the Collaborative Transportation is the preparation of a shipping prevision, including the generation of orders and loads and, finally, the performance of the delivery and payment to the carrier (Silva *et al.*, 2009). The collaboration must act in the formation of groups (with the same agents: industries, per example) to reinforce power to do the negotiations with the other group of agents (per example, the shipowners).

For the implementation of the CTM it is required, therefore, systems allowing inter-organizational collaboration in a cost-effective and technologically compatible manner. Without such systems, the companies trying to implement the systematic may find invalid efforts which are difficult to manage.

3.1 Collaborative transportation operational system models

The general process of decision taking begins with the finding of the decision taking problem, which means, a challenge for future business operations such as long term operations in the transportation market when entering collaborative system (Bloos and Kopfer, 2009). According to these authors entering in collaboration operational system in transportation implies in the exchange of client data, which can be a very sensitive operation. The process begins with the searching for alternative solutions and an accurate specification shall be performed on how the goal can be achieved and measured. After reaching, the costs and benefits are distributed during the transactions.

Gomber *et al.* (1997) present a collaboration model for planning of transportation through the usage of freight agencies with several autonomous rewarding centers actuating in the achievement of orders and negotiation of prices for the execution of the orders of the clients. Besides the short literature on this subject, some works were found specifying operational systems for collaborative transportation, which may serve as benchmarking for the companies: Kopfer and Pankratz (1999), Grünig and Kuhn (2005), Schönberger (2005), Krajewska and Kopfer (2006), Bloos and Kopfer (2009).

Regarding the applied part, involving the construction of mathematical model and application, one of the first Brazilian works referent to the collaborative transportation subject is of Carnieri *et al.* (1983). It analyses the feasibility of the implementation of a representative central of the cooperatives in the State of Paraná (Brazil), aiming at minimizing the cost of soy transportation. The problem is modeled as Integer Programming.

Tacla (2003) develops a methodology for collaborative transportation of soy and fertilizer introducing flexibility to the restrictions imposed by time windows and minimizing freight expenses. Using mathematics model and CPLEX solver of GAMS 2.50® and, for the operational level, develops a heuristic solved in Excel® suggesting Linear Programming problem. Different from the traditional approach, Novaes *et al.* (2009) consider in its analysis a dynamic environment for urban routing problem subject to traffic jams. Using the concepts of Sequential Analysis, the study comprises a warehouse and a homogeneous fleet of vehicles serving a certain urban region.

As observed, there are different approaches for the transportation problem, involving several sources, destinations, companies and products. A possible manner to model the maritime transportation problem for the Brazilian exportation of manufactured products is by the recent application of Autonomous Agents. In the following item it is presented the concept of Autonomous Agents applied to logistic problems.

4. APPLICATION OF THE AUTONOMOUS AGENTS CONCEPT IN THE RESOLUTION OF TRANSPORT LOGISTICS PROBLEMS

The most part of data process systems and logistics systems comprise several components, treated as agents, which usually actuates in an autonomous manner, which can also communicate and interact with the others, forming a network. Therefore, the Autonomous Agents form a community in a common environment where they actuate and transform based on rules, in order to reach their individual objectives. Besides being a new approach to logistics, several studies have been developed by the Collaborative Research Centre 637 Autonomous Cooperating Logistic Processes group in Bremen, Germany, with the purpose of detecting when and how the autonomous control can present more advantages than the classic control, especially regarding time, costs and robustness.

4.1 Autonomous agent concept

Agents are components in the complex adaptative systems helping on decision-making. According to North and Macal (2007) agents have sets of rules of behavior patterns that allow them to take in information, process the inputs and then effect changes in the outside environment. Adaptation and learning are the results obtained by the agents after processing the information between them. More specifically, these authors define agents like anything that makes choices in a business: managers, executives, organizations.

In Agent Based Modeling Simulation (ABMS), an agent is an individual with a set of attributes and behavioral characteristics. The attributes define what a given agent is. See the example below in Figure 2.

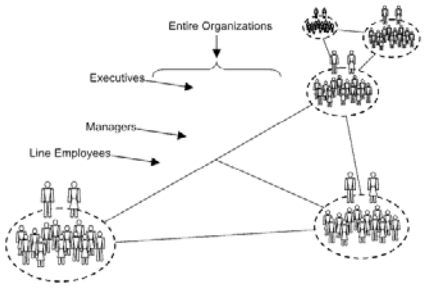


Figure 2. Example of real-world agents
Source: North and Macal (2007)

The behavioral characteristics define what a given agent does and, generally there is difference among many of the agents in the model. To start modeling an agent is necessary to define its attributes, which can vary from one to another. Some common attributes to represent people include age, income, sex, history; to represent corporations competing within markets include resources, size, decision time and various strategic preferences such as risk tolerance. The next step to continue modeling is to define the agents' behaviors. Common behaviors include operations and planning. Some features include decision rules to select actions, adaptation capabilities to learn from experiences, perceptual capabilities to sense its surroundings, and optional internal models to project the possible consequences of the decisions.

Casti (1998) defines essentially two levels of agent rules. The first are agent-based rules. These rules specify how the agent responds to routine events. The second contains "rules to change the based-agent rules", which provide adaptation by allowing the routine responses to change over time. Agents have sets of decision rules that govern their behaviors. These rules allow agents to interact with and communicate with other agents as well as to respond to their environments. These rules can provide agents with responsive capabilities on a variety of levels from simple reactions to complex decision-making. The first step followed by the agents is the evaluation of their current moment. The second step is the execution of the actions that they have chosen and the third step is the evaluation of the results of their actions and the adjustment of their rules based on the results. These steps can be performed with the use of simple rules, complex rules, advanced techniques, external programs, or even nested subagents.

Despite being a new concept applied to logistic problems, researches advocate this tool gives a decentralized control for the supply chain, including local information processing, where the agents can step by step take a good decision. It is presented in the following item a model prototype for the resolution of the transportation problem of this work, considering the agents as autonomous actuating in a dynamic environment.

5. PROTOTYPE MODEL OF RESOLUTION OF THE MARITIME TRANSPORTATION LOGISTIC PROBLEM

Due to the fact that the matter of the study is very comprehensive, with different variables and agents involved, it shall not be handled as an enhancement problem. Therefore, the proposed model has the objective of establishing scenarios involving the main agents existing in the maritime transportation for better analyzing the dynamic behavior in the negotiation. Such agents are: industries, third-part logistics provider, NVOCC, shipowner. It must consider intrinsic information of all these agents like due date, deadlines, costs of transportation, port fees on load moving, etc., in order to reduce the cost of the collaborative chain, or the time, if it is under its interest.

In the development of the model the behavior of the environment shall be judged as dynamic, with random transportation demands and requiring the allocation of the manufactured products in containers. Both participants in the maritime logistic shall be judged as autonomous agents, observing the imposed restrictions and taking decisions about appropriate behavior between the origin of the load and its final destination, in order to achieve a feasible solution for the logistic network.

Another matter to be highlighted is the set of advantages offered to the participants of the collaborative chain: in order to achieve alterations in the shipping and delivery dates to its clients, it shall achieve gains with the reduction of freight resulting from the joining of loads; which means, the costs of the collaboration shall be lower than the individual actuation costs. Therefore, the proposed model shall comprise a systematic of distribution of these gains in order to make attractive the participation of the entrepreneurs. A possible way to distribute these costs is using the *Shapley value* which is better described on Shapley (1953).

5.1 Model details

In the Figure 3 it is possible to find some inputs and outputs of the model for the performance of maritime transportation:

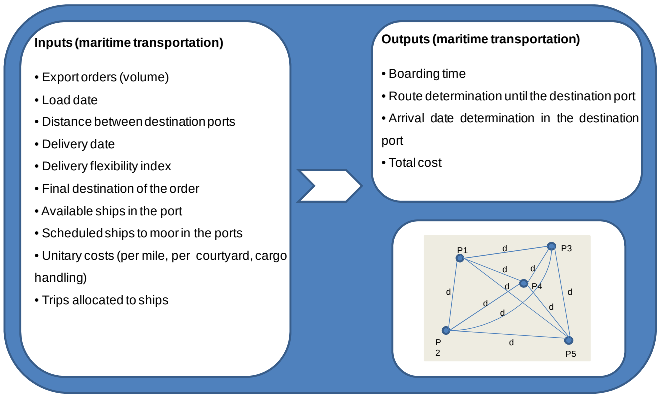


Figure 3. Input and output information of the maritime transportation model
Source: author.

It must be known the volume cargo by the industries, storage costs, loading/unloading costs, freight, benefits and injuries to the partners. The same must happen one step before, for the land transportation completing the maritime network.

It can be established a description of the model initially through the overview of the process presented below in six stages, completing with the flowchart presented in

Figure 4:

a) The companies, in an independent and dissociated manner, executing its respective business plans comprise the sale of its manufactured products. From this plan the logistics planning areas shall create a storage and transportation plan. In this stage, the proposed model interacts with the individual plans creating a collaborative transportation planning, initially until the source port, considering shipping and delivery dates in the destination, as well the definition of the volumes. The collaborative transportation is initiated in this stage.
b) To perform the land transportation between the collaborative companies until the source port.
c) Verify the possibility of prompt shipping in the ship or waiting in the port apron.
d) The ship verifies if there is load for shipping or landing in the ports of its route and chooses to berth in it or not.
e) After shipping, the container can follow in the same ship directly to the destination port or choose for the landing in some port of the route to achieve better arrival date in the destination or better cost, considering the possibility of advance or prorogation in the delivery of the goods. The collaboration in this stage reaches not only the operational planning levels, but also the areas directly linked to the business. The Sales departments of the exporter companies shall agree in advancing or delaying deliveries to its clients, opting for making more flexible or not the time windows.
f) The transportation programming shall be the performance of the transportation of the exportation orders, aiming at the minimization of the costs and compliance with dates, closing the cycle.

With the analysis of the several scenarios varying the volumes, flexibility date, costs, among other factors on the model it is expected to observe the behavior of the maritime network when working on collaboration between its agents. The purpose is to identify situations which lead to scale savings transferred to collaborators in order to enhance the efficiency of the logistics and the financial result of the agents, how actually proposes the collaborative transportation approach.

The freight reduction can be obtained when using the same unit of transportation or better negotiating prices with the shipowners by the industries groups formation. The problem can be judged as a simple event in a first moment, but actually the combination of several possibilities between the collaborators of the system makes its execution very complex. Therefore, to achieve success in the resolution of this collaborative transportation problem, it is critical the implementation of an structured mathematical model in order that the system presents a feasible and accurate solution, with a total final cost lower than the cost obtained with the routes composed by the individual companies.

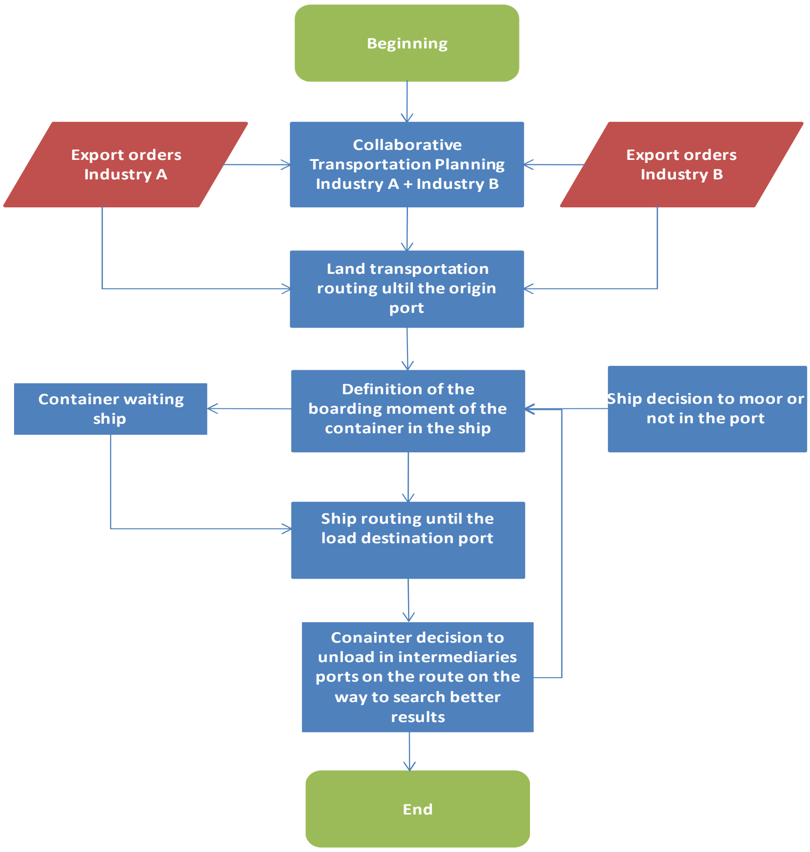


Figure 4. General flowchart of the proposed model
Source: author.

6. FINAL CONSIDERATION

Among the several links existent in the composition of the logistics chain, in can be observed that the transportation activity is one of the most significant of the chain, once it is through transportation that it is possible to synchronize the availability of raw material on time, in full in the manufacturing units and the final products in the distribution centers further to the final client (Silva, 2008). In this context, therefore, the CTM has been used in order to allow the synchronization of the logistics network.

Therefore, considering the relevance of this tool and there is demand for the maritime transportation service by the exporter companies, it shall require researching and modeling the transportation problem discussed hereby, analyzing the possible scenarios which best understand the needs of the companies, in order to search the equilibrium between its gains and losses. As presented hereby, the Autonomous Agents approach applied to this type of problem may constitute a tool to analyze the actuation scenarios of the exporter companies requiring discharging its production through maritime transportation.

Considering the lacking of studies in this area it is expected that this study may be useful as incipient step in the development of a supporting tool for the analysis and decision takings by the entrepreneurs. Therefore, for the continuity of the study it is required to calculate the model on a mathematical basis, develop decision criteria to distribute the costs and benefits among the main agents involved in the collaborative transportation and test them. From the calculated model it is recommended to define one useful software to simulate this transportation problem based on the Autonomous Agents theory interacting in dynamic systems. After that, it is recommended to simulate operations in different operational conditions similar to the corporate practices in order to test the model and validate it.

ACKNOWLEDGEMENTS

This research h as b een supported b y th e Brazilian CNPq /Fapesc Pro ject n º 1.0810-00684 (Pron ex), b y Capes/DFG Bragecrim Project nº 2 and by the Bremer Instituts für Produktion und Logistik (*BIBA*) an der Universität *Bremen.*

7. REFERENCES

Bloos, Melanie; Kopfer, Herbert. (2009). On the formation of operational transport collaboration systems. *Annals of the II LogDynamics International Conference (LDIC)*: 329-338. Bremen, Alemanha, 17-21 August.

Carnieri, Celso; Simiema, Hélio Hipólito; Mazzarotto, Marco André (1983). Programa Integrado de Transporte de Soja. *Anais do XVI Congresso da SBPO,* Florianópolis.

Casti J. (1997). *Would-Be Worlds: How Simulation Is Changing the World of Science.* New York: Wiley.

Gomber, P; Schmidt, C ; W inhardt, C . (1 997). Elektronische M ärkte für die deze ntrale Transportplanung. *Wirstschaftsinformatik,* 39 (2):137-145.

Grünig, R ; Kühn, R . (2005). *Entscheidungsverfahren für komplexe Probleme. Ein heuristischer Ansatz.* 2nd ed ition, Berlin, Heidelberg, New York: Springer.

Hölscher, Ka rsten; Kl empien-Hinrichs, R enate; K nirsch, Pet er; K reowski, Hans-Jorg; K uske, Sabine. (2007). *Autonomous units: basic concepts and semantic foundation*, In: Hülsmann, M.; Windt, Katja (ed.) Understanding autonomous cooperation and control in logistics, German: Springer.

Keedi, Samir. (2007). *Logística de transporte internacional.* 3ª ed. São Paulo: Aduaneiras.

Kopfer, H; Pan kratz G . (2009). *Das groupage-problem koopererender Verkehrsträger.* Op er. R es. P roc., 1 998: 453-462, Springer, Berlin Heidelber New York, 1999.

Krajewska, Marta Anna; Kopfer, Herbert. (2006). Collaborating freight forwarding enterprises – request allocation and profit sharing. *OR Spectrum* 28:301-317.

Mdic – Ministério do Desenvolvimento, Indústria e Comércio Exterior website.

North, Michael J; Macal, Charles M. (2007). Managing business complexity: discovering strategic solutions with agent-based modeling and simulation. New York: Oxford University Press,.

Novaes, An tonio G . N; Frazzo n, En zo M; Burin, Pau lo J. (2009). *Dynamic* vehicle ro uting i n o ver congested u rban areas. *Annals of the II LogDynamics International Conference (LDIC)*: 103-112. Bremen, A lemanha, 1 7-21 de agosto.

Schönberger, J. (2005). *Operational freight carrier planning*, Berlin, Heidelberg, New York: Springer.

Shapley, L. S. A. (1953) V alue for n-Person Games. In : KUHN, H . W.; TUCKER, A . W. (E ds.) Co ntribution t o the Theory of Games. *Annals of Mathematics Studies*. New Jersey: Princeton University Press, p. 307-317.

Silva, Vani na M acowski Du rski. (2 008) *Um modelo heurístico para alocação de navios em berços.* Dissertaçã o de mestrado do Programa de Pós Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina, Florianópolis.

Silva, Va nina M acowski D urski. Zago, C amila. A. A breu, Leonor. F. Coelho, Antonio. S. G onçalves, Mirian B. (2009). La i nfluencia de l a gestión del t ransporte co laborativo para la eficien cia d e la cadena d e suministros. *Anales del XIV International Conference on Industrial Engineering and Operations Management (ICIEOM)*, Salvador-BA, Brasil.

Souza, E rika. (2009). Al ianças, América L atina e si stema m arítimo po rtuário m undial. *Anais do* 12^0 *Encuentro de geografos de America Latina*, 3-7 de abril, Montevideo, Uruguay.

Tacla, Douglas. (2003). *Estudo de transporte colaborativo de cargas de grande volume, com aplicação em caso de soja e fertilizantes.* Tese de Doutorado. Programa de Engenharia Naval e Oceânica, área de concentração: Transportes e Sistemas Logísticos, Escola Politécnica da Universidade de São Paulo.

DEVELOPMENT OF OPTIMIZATION MODEL FOR SERVICE SUPPLY CHAIN

Min Kyu Hwang[1] and Young Hae Lee[2*],

Department of Industrial and Management Engineering
Hanyang University
S. Korea
[1] munggu@hanyang.ac.kr
[2*] yhlee@hanyang.ac.kr

Abstract: The importance of service industry is increased continuously, and there is a change from product-domination to service-domination due to appearing of service science. Thus, needs for understanding and studying service system are increased. Recently, perception and attention about the importance of service have been increased, but the research for them is weak. Especially, the SCM research which was perceived as an essential factor for achievement of competitive advantages is based on the manufacturing industry, so there is a limitation for applying to the service industry. In this article, the service supply chain is analyzed systematically based on service science, and the model for optimization of operation is set up. Also, the framework for using the simulation optimization technique is established, and there is a detailed example of the optimization model of service supply chain which was developed in the service industry.

Keywords: service; service supply chain; service science; optimization

1. INTRODUCTION

The importance of service industries has been increasing because of accelerating the transition to knowledge-based economy. In the past, service industry has been recognized as a secondary of manufacturing industry. Now the service industry is becoming new industries to drive the economy.

Service industry is already a huge part in the world economy and getting bigger. Because people recognized this situation, the topic which is service science and service is researched and interested from many people. But the concept, theory and application field etc. aren't completed. The research on service industry is continuing because of the necessity of service.

Manufacturing industry which is focused on products than service is changing by time flow. We can easily find the example of this case. For instance, water purifier industry sales the water purifier generally. However, the traditional structure is changed to different sales structure. They rent the water purifier and provide steady management of the machine. So the sale is rent fee. Like this way, companies can satisfy variety needs of customers. Also, companies can keep relationship with customers as well as sale their service. This chance is a new strategy of companies to survive.

The past research which is relative with supply chain is almost focused for manufacturing industry. It can be used to service industry. Because the industry structure is changing from manufacturing industry to service industry, the research on service industry is needed.

Especially, the SCM research which was perceived as an essential factor for achievement of competitive advantages is based on the manufacturing industry, so there is a limitation for applying to the service industry.

The Purpose of this research is to understand unique characteristics of service and to make a new optimization model through applying the characteristics to service supply chain. There is a detail example of the optimization model of service supply chain using the developed service model.

2. SERVICE SUPPLY CHAIN

2.1 Concept of service science

Necessity of service science is needed for a long time but the academic research is weak. Service science is to redefine the fundamental service and develop the service level. Also, that is to increase productivity through business administration, social science, industrial engineering etc.

---

[*] Corresponding Author

2.2 Concept of service supply chain

The serv ice SCM d efinition is, "As d escribed earlier, Ellran *et al.* (2004) a nd B altacioglu *et al.* (2007) d eveloped service SCM model, but their service SCM model is limited to general characteristics of service. Also, it is focused on delivering service in SCM and not considered intersection between the members. In addition, there is no considering about elem ents of affecting to value creation. This can be solved by recent service science c oncept which is new recognition of service.

2.3 Service entity and relationship

Service entities consist of service resource supplier, service provider, service customer. Each entity has different role in different situation. Figure 1 is representative relationship between entities.

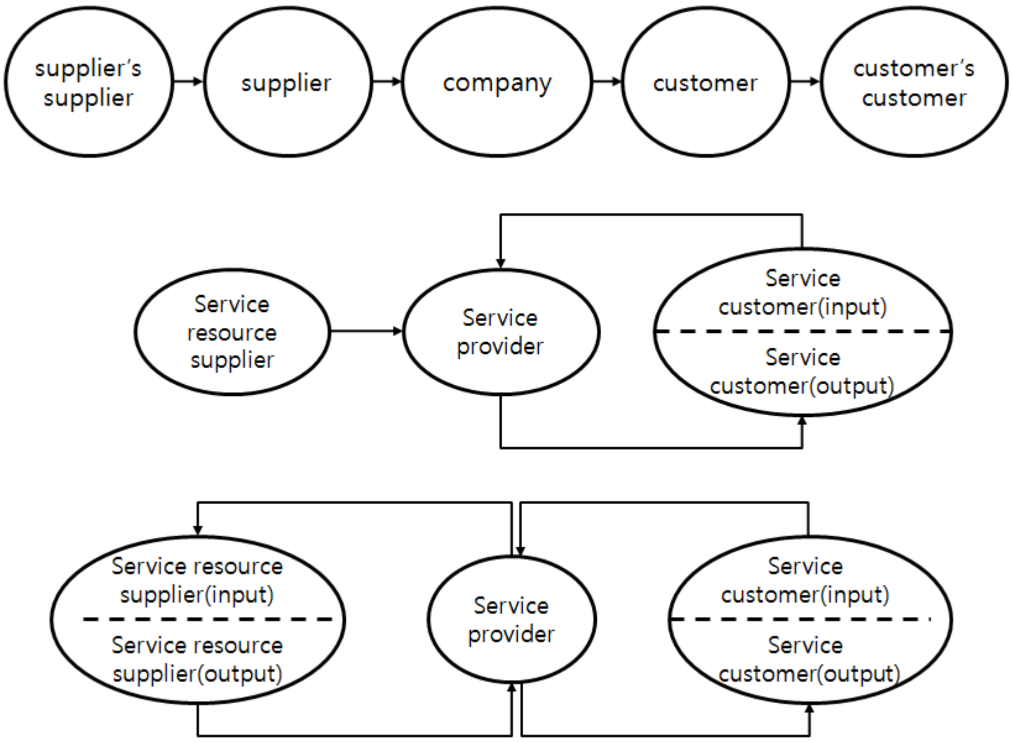


Figure 1. Relationship of each service entities

In Figure 1, the top chain is common supply chain of manufacturing industry. The flow is from under supplier to upper customer through the company. The arrow means product flow. The second chain is bidirectional supply chain model of level 1. Service custom er uses the output and ret urns t he input t o serv ice prov ider. The last ch ain has two level relationships. The service resource supplier receives service of service customer from se rvice provider and gives the output of re-product to service provider. The second chain and the third chain have the arrows which are product flow, too. The dotted lines are divided by service process.

3. APPLICATION OF SERVICE SUPPLY CHAIN MODEL

3.1 Framework of service system

There is a process of virtual mobile phone re pair ce nter through the rela tionship of each se rvice entities and t he developed mathematical model. We use the process of virtual mobile phone repair center to apply our developed model. This framework of service system will be used for the base of service simulation model.

3.2 Application

The mobile phone repair center starts all process when the customer arrives. Arriving customer is considered as input resource and leaving output which is completed in a process is considered as repaired mobile phone. There are two repairmen in the repair center. Each repairman takes different time to same work. The same work is same price, though different repairmen did it. This can be a tree structure in Figure 2.

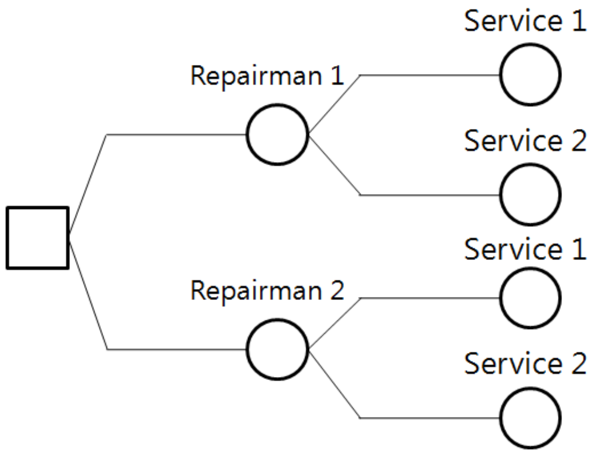


Figure 2. A tree structure of mobile phone repair center

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0016024)

4. REFEREBCES

Ahn S. H. and Lee Y. H. (2009). Service Science and Service SCM. Korean Society of Supply Chain Management Spring Conference.

Arash S. (2010). SSCM: Service Supply Chain Management. *International Journal of Logistics Systems and Management*, 6, 1: 60-75.

Baltacioglu, T., Ada, E., Kaplan, M. D., Yurt, O., and Kaplan, Y. C. (2007). A New Framework for Service Supply Chains. *The Service Industries Journal*, 27, 2: 105-124.

Chopra, S. and Meindl, P. (2007). *Supply Chain Management*. Pearson Education.

Li, C., Liu, Y. and Cheng, J. (2008). The Research on Service Supply Chain. IEEE International Conference on Service Operations and Logistics and Informatics, 2 Digital Object Identifier: 10.1109/SOLI.

Cook, J. S., DeBree, K., and Feroleto, A. (2002). From Raw Materials to Customers: Supply Chain Management in the Service Industry. *SAM Advanced Management Journal*, 66, 4: 14-21.

Ellram, L. M., Tate, W. L., and Billington, C. (2004). Understanding and Managing the Services Supply Chain. *Journal of Supply Chain Management*, 40, 4: 17-32.

Fitzsimmons, J. A., and Fitzsimmons, M. J. (2007). *Service Management*, McGraw-Hill.

Georgantzas, N. C. (2003). Tourism Dynamics: Cyprus' Hotel Value Chain and Profitability. *System Dynamics Review*, 19, 3: 175-212.

Jones, P. (1999). Operational Issues and Trends in the Hospitality Industry. *International Journal of Hospitality Management*, 18, 4: 427-442.

Kathawala, Y., and Abdou, K. (2003). Supply Chain Evaluation in the Service Industry: A Framework Development Compared to Manufacturing. *Managerial Auditing Journal*, 18, 2: 140-149.

Korpela, J., Lehmusvaara, A., and Tuominen, M. (2001). Customer Service Based Design of the Supply Chain. *International Journal of Production Economics*, 69, 2: 193-204.

Maglio, P. P., Srinivasan, S., Kreulen, J. T., and Spohrer, J. (2006). Service System, Service Scientists, SSME, and Innovation. *Communications of the ACM*, 49, 7: 81-85.

Maglio, P. P. and Spohrer, J. (2007). Fundamental of Service Science. *Academy of Marketing Science*.

Nam, K. C., Kim, Y. J., Nam, J. T., Bae, Y. W., Byun, H. S., and Lee, N. H. (2008). Service Science: Theory Review and Development of Analytical Framework. *Information Systems Review*, 10, 1: 213-235.

Service Science Vortal (2006). *Service Science*, Maeil Business Newspaper.

Simchi-Levi, D., Kaminsky, P., and Simchi-levi, E. (2008). *Designing and Managing the Supply Chain: Concepts, Strategies, and Case Studies*. Mcgraw Hill.

Veronneau, S. and Roy, J. (2009). Global Service Supply Chains: An Empirical Study of Current Practices and Challenges of a Cruise Line Corporation. *Tourism Management*, 30, 1: 128-139.

Wang Chun-xi (2003). Research Status and Developing Trends of the Performance Measurement of Supply Chain. *Journal of Industrial Engineering and Engineering Management*, 17, 3: 27-29.

Session G2: Intermodal

·Day2: Sep. 16 (Thu.)

·Time: 13:00 - 14:20

·Chair: Takashi Amemiya

·Room: Azalea & Lilac, 5F

A HEURISTIC ALGORITHM FOR MULTI-PERIOD, MULTI-MODAL TRANSPORTATION PROBLEM

Ikju Jang* **and Suk-Chul Rim**

Department of Industrial Engineering
AjouUniversity
San 5, Wowncheon-Dong, Yeongtong-Gu
Suwon, Korea
e-mail: bk99@ajou.ac.kr

Abstract: Today's global companies are executing global supply chain management which delivers huge amount of products to the customers rapidly via global transportation network that connects factories, distribution centers, and customers located around the world. In most companies, since the transportation cost exceeds 60% of the total logistics cost, it has to be reduced to reduce the total logistics cost. Unit transportation cost varies from the fastest but most expensive air-parcels to land (truck), and the slow but cheapest sea (vessel) route. The combination of multiple transportation modes are often used to transport goods to their destination, namely, multi-modal transportation.Existing studies on the multi-modal transportation mostly have dealt with the static problem in which to transport goods from an origin to a destination with minimum total transportation cost. However, in practice, companies have to make decisions periodically as to how much to send via which route to minimize the total cost. So the problem in practice is a dynamic or multi-period multi-modal transportation problem.In this paper we study a multi-period, multi-modal transportation problem for a single product,where we make decision every week as to how much to send to the customer via which route, considering the current inventory and demand forecast at the customer site so as to maintain a certain service level and minimize the inventory, thereby minimizing the total cost, which is the sum of the transportation cost and holding/obsolescence/shortage cost. We propose a practical heuristic algorithm for this problem. Simulation is used to evaluate the average performance of the proposed algorithm.

1. INTRODUCTION

1.1 Problem Description

In today's global business, many manufacturers operate global supply chain to transport products via transportation network that connects many manufacturing sites, distribution centers, and customer sites located around the world in short time. Especially for high-tech electronic products such as mobile phone, personal digital assistants, or LCD TV, due to rapid technological changes and short life cycle, the market is very hard to forecast. Furthermore, many of these products are produced at the sites such as China and Korea that are far away from the major markets such as US and Europe. Therefore meeting volatile demand at the customer site at the lowest transportation and inventory cost is a critical issue of running supply chain in today's highly competitive business environment.

In most global manufacturers, since transportation cost exceeds two thirds of the total logistics cost, it is very important to reduce the transportation cost in order to reduce the logistics cost. For global transportation, multiple transportation modes are available, which usually include air freight, vessel, rail-road, and truck. Each mode has its own advantages and disadvantages. In general, faster mode is more expensive for unit amount of product. So there is always a trade-off between cost and speed in choosing the transportation mode. Unlike the static problem as many optimization problem assume, the practical situation is that a company needs to make decision dynamically as to how many products to ship to its destination at every period of time such as week.

In this paper, we define the dynamic multi-modal transportation problem(DMTP) as the problem of determining at every period the shipping amount of a product from an origin to a destination via multiple transportation mode with respective lead time and unit cost so as to minimize the total transportation and inventory cost, while maintaining the customer service at the destination to a certain level.The difficulty of running the DMTP is that the certain portion (e.g., 95%) of demand must be met by on-hand inventory at the destination, while some transportation modes take longer lead time than the length of decision period. So DMTP can be viewed as a variation of the well-known periodic review inventory control model (R,s,Q). The difference is that (R,s,Q) automatically determines the order quantity as (base stock - current on-hand inventory); while DMTP has to determine the order quantity (or shipping amount, from the origin's viewpoint) for each of multiple transportation modes. For objective function, (R,s,Q) is designed to maintain the specific service level; while DMTP aims to minimize the total cost and maintain a specific service level.Consider two extreme choices; the most expensive air transportation which will enable the inventory level to be reduced to minimum, and the least expensive one by vessel which will result in higher inventory level to cope with possible rapid increase of demand during the long lead time.

The relative economic advantage of the two alternatives will depend on the parameter values, which can be classified into product parameters, cost parameters and modal parameters. Product parameters include price of unit product, profit rate, and available quantity from the origin at every period. Cost parameters include inventory holding cost, shortage cost, backorder cost, and obsolescence cost which is significant for high-tech products. Modal parameters include unit transportation cost and lead time. Given these parameter values, there will be an optimal choice of shipping quantity via each mode between the above two extreme choices. In this paper, we will propose a heuristic approach for the problem.

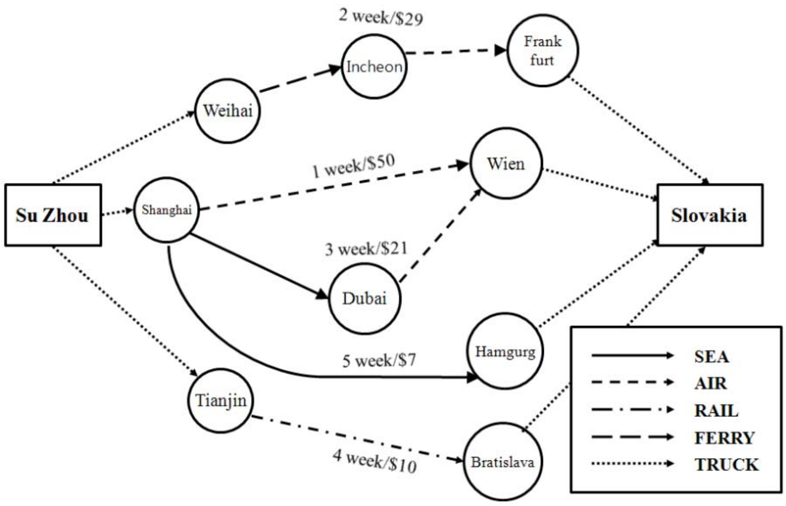


Figure 1.An illustrative example of Dynamic Multimodal Transportation Problem with 5 routes from Su Zhou to Slovakia.

To illustrate the DMTP, consider the following example in Figure 1. From Su Zhou, China to Slovakia,the product can be transported via five different routes, each of which is a combination of five different transportationmodes of sea, air, rail, ferry, and truck. Lead time varies from one week to five weeks; and unit transportation cost varies from $7 to $50.

If the assembly plant at Slovakia operates the make-to-stock (MTS) production, the problem is simple; the lowest cost ordering is to order the required amount from master production schedule (MPS) via the least expensive mode ahead of the lead time of the mode. However, for most high-tech electronic products, make-to-order (MTO) must be employed in today's volatile market due to its short life cycle and relatively high price.

Every business unit forecast demand by aggregating all available sales information. Although the forecast accuracy is higher as time approaches to the period to forecast, forecast accuracy cannot be guaranteed due to many factors such as competing product, technological changes, marketing and promotion, especially for high-tech electronics.

In this paper we address the multi-period dynamic problem of determining at the beginning of every period the shipping quantity of single product via multiple routes by considering the on-hand inventory at the destination, in-transit inventory, and demand just realized at the end of the previous period at the destination, so as to minimize the total cost and maintain the required service level at the destination.

The contribution of this study can be substantial. Consider a global manufacturer with annual sales of $100 billion. Suppose the logistics cost is 10% of the total sales; profit rate is 3%, export logistics cost is 70% of the total logistics cost, 70% of which is export transportation cost. Suppose the company successfully operates the multi-modal transportation in their export business, and save the export transportation cost by 30%. Then, out of $10B of total logistics cost; the export transportation cost saving of $1.47Billion will result in 49% increase of the total profit, which is substantial; and provide higher competitiveness in global market, especially for high-end products with short product life cycle.

1.2Assumptions

We assume the followings in this paper.

- Decision is made at the beginning of every period (without loss of generality, we assume a week as a period.)

- Transportation cost is proportional to the shipping quantity for any route.
- All routes always have transportation capacity available as much as desired.
- Lead time and unit transportation cost is constant for each route.
- Product is always available as much as desired.
- Demands during the periods are independent, and identically distributed random variables; and follow a normal distribution.

2. LITERATURE REVIEW

Most research results on determining shipping quantity over multiple routes use updated demand forecast; and based on the demand forecast method. Sethi*et al.*(2001) classify these into the following three categories. The first category is Bayesian analysis, where demand forecast is updated by Bayes rule (e.g., Eppen and Iyer, 1997). The second category is using time-series models, which is useful for the cases when demands of consecutive periods are correlated. Demand can be modeled as an ARIMA model; and myopic optimal policy can be derived under special conditions. (e.g., Johnson and Thompson, 1975; and Lovejoy, 1990)

The third category is demand forecast update model. Hausmann(1969) present an optimal decision rule by modeling the updating process of demand forecast as time elapses as the quasi-Markovian process. Sethi and Sorger(1991) propose an optimal solution framework for rolling-horizon type decision rule for practical use. However, their dynamic programming formulation has a limited practicality due to heavy computational burden. Heath and Jackson(1994) model the evolution process of demand forecast using a martingale; and address the problem of determining economic safety stock level for multi-item, multi-facility production system. Yan *et al.*(1998) present a single-period optimal ordering rule for two supply channels case with updated demand forecast.

Donohue(1998) study the pricing policy for perfect demand update in the risk-sharing purchase contract between seller and buyer.Barnes-Schuster *et al.*(1999) study the structure of objective functions between seller and buyer in single period, two-stage model where updated demand forecast arrives at the beginning of the second stage.Gallego and Ozer(1999) model the updating process of demand forecast using super martingale; and show that the state-dependent (*s,S*) policy is the optimal. Bensoussan*et al.*(1983) study the inventory model with two supply modes, one instant delivery and the other with one unit of lead time; and present generalized (*s,S*) as an optimal ordering policy for fixed or variable ordering cost.

Hausmann*et al.*(1993) present an optimal ordering policy for the inventory model with two supply modes of fast and slow delivery under stationary demand. Zhang(1996) extends this study to the case of three supply modes. Scheller-Wolf and Tayur(1998) show that state-dependent base-stock policy is the optimal for Markovian dual-source production inventory model without taking demand forecast into consideration.

Sethi*et al.*(2001) address the problem in which demand forecast is updated as time elapses. The authors point out that future demand contains many randomness factors; and uncertainty is eliminated as time approaches, finally real demand is realized, just as peeling off an onion to the core. This study deals with the multi-period problem of periodic review ordering for single item; decision variable is shipping quantity for fast and slow route at every review time. Objective function is to minimize the total cost comprising ordering cost and inventory holding cost, including backlog cost. The authors show that the ordering policy defined by the state-dependent base-stock level of fast and slow route is optimal for finite-period and discounted infinite-period problems. As an extension by the same authors, Sethi*etal.*(2003) show the optimality of (*s,S*) policy by demand forecast update when ordering cost is constant.

Kim(2009) propose a heuristic algorithm to determine the shipping quantity of single product for multiple route, in order to avoid shortage using periodically updated demand forecast. However, this study employs some vague concept such as 'week of supply' which is not well defined in academic viewpoint. For the papers that address the static version of the multi-modal transportation problem, the readers may refer to Chang(2008). Song(2006) considers a multi-period production, inventory and distribution planning issues using Benders' decomposition approach where distribution issue is represented as a dynamic assignment model.

Our study is unique in that, unlike the most of the papers introduced above, our study does not consider demand forecast in determining the shipping quantity over multiple routes, but assume demand follows certain distribution. This is based on the interviews with practitioners that the demand of high-end electronics with short life cycle is so hard to forecast that it will be more appropriate to regard demand as a random variable that follows a certain probability distribution.

3. HEURISTIC ALGORITHM

3.1Modeling

We will use the following notations to model the DMTP.

S_t^i: shipping quantity in route i at the beginning of period t. (decision variable)
α : probability of shortage in a period.
i: route (i=1, 2, 3, …, N)
t: period(t=1, 2, 3, …, T)
D_t : demand during period t.(realized at the end of period t)
E_t : ending on-hand inventory at end of period t.
A_t : arriving quantity at beginning of period t.(or, equivalent, at end of period t-1)
O_t : on-hand inventory at beginning of period t.($O_t = E_{t\text{-}1} + A_t$)
c_i : unit transportation cost of route i.(c_1=\$50, c_3=\$21)
B_t : backorder quantity at end of period t.(B_t=max[0, - E_t])
P: price of the product (\$600)
b: backorder cost (10% of P)
h: holding cost (25% of P per year)
ψ : obsolescence cost (40% of P per year)

To simplify the problem, without loss of generality, let us consider only two routes whose lead times are one and three weeks, respectively. A$_t$ the beginning of period t, whenS^1_{t-1} and S^3_{t-3} have arrived, we need to determineS^1_t and S^3_t, taking into account the two in-transit inventory S^3_{t-1} and S^3_{t-2} , as shown in Figure 2.

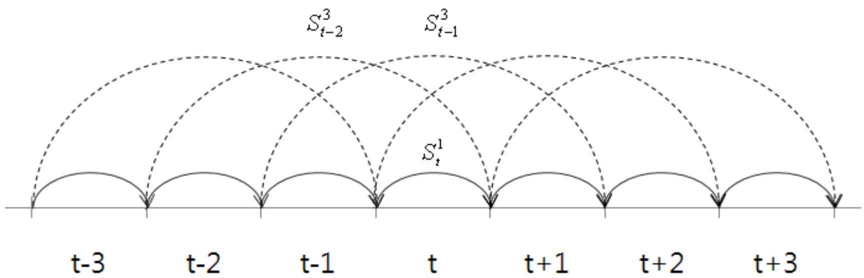


Figure 2.Two in-transit inventory at beginning of period t.

One way of determining the shipping quantity $S^1_t$and S^3_tat beginning of period t is that we assign a certain portion (e.g., 80%, or λ=0.8) of the average demand toS^3_t ;and determine S^1_tsuch that the desired service level is attained. We assume stationary demand; that is, D_t follows the normal distribution with mean μ, and standard deviation σ. We need to determine S^1_tsuch that the demand in period (t+1) will have shortage with probability α. Note that S^1_twill arrive at the destination at end of period t; and be available for the demand in period (t+1).

By definition, for 1≤t≤T,

$$\Pr(D_t > E_{t-1} + A_t) = \alpha \quad (1)$$
$$\mathrm{E}_t = \mathrm{E}_{t-1} + \mathrm{A}_t - \mathrm{D}_t \quad (2)$$
$$\mathrm{A}_t = \mathrm{S}_{t-3}^3 + \mathrm{S}_{t-1}^1 \quad (3)$$

Using (2) and (3), equation (1) can be rewritten as

$$\Pr(D_{t+1} > E_{t-1} + A_t - \mathrm{D_t} + \mathrm{S}_{t-2}^3 + \mathrm{S}_t^1) = \alpha \quad (4)$$

So, $\Pr(D_t + D_{t+1} > k) = \alpha, where\ k = E_{t-1} + A_t + S_{t-2}^3 + S_t^1$ (5)

Since D_t are assumed to be iid, $D_t + D_{t+1} \sim \mathrm{N}(2\mu, 2\sigma^2)$; so k can be computed as (6).

$$k = 2\mu + z_\alpha \sigma\sqrt{2} \tag{6}$$

Once k is obtained from (6), then the one-week shipping quantity S^1_t can be determinedfrom(5) and (6) at beginning of every period t.

$$S^1_t = k - (E_{t-1} + A_t + S^3_{t-2}) \tag{7}$$

Now we define the total cost incurred during the period $t$as the sum of transportation cost, inventory cost, obsolescence cost, and backorder cost over all routes, as (8), where IP_t denotes the inventory position at beginning of period t, as defined in (9).

$$\text{Total cost} = (c_1 \cdot S^1_t + c_3 \cdot S^3_t) + IP_t \cdot (\mathrm{h} + \psi) + B_t \cdot b \tag{8}$$
$$IP_t = O_t + S^3_{t-1} + S^3_{t-2} - B_{t-1} \tag{9}$$

We assume that backorder is cleared as soon as the products become available at their arrival at beginning of the period.

Table 1. Computed data for the numerical example

t	S^3_t	A_t	O_t	S^1_t	D_t	E_t
1	80	0	100	123	111	-11
2	80	123	112	111	95	17
3	80	111	128	15	82	46
4	80	95	141	2	92	49
5	80	82	131	12	109	22
6	80	92	114	29	102	12
7	80	109	121	22	109	12
8	80	102	114	29	121	-7
9	80	109	102	41	98	4
10	80	121	125	18	98	27

3.2Numerical example

Consider an examplewhere $D_t \sim N(100, 10^2)$, initial inventory O_1=100, α=0.95, λ=0.8. Since $z_{.95}$=1.645, the constant value of k can be obtained from (6) as k=223. Then the values of S^1_tcan be computed using (7), as shown in Table 1for t=1, ...,10.Obviously we can confirm that the values of S^1_tfollow a distribution with mean=20 except for the first two periods during which $S^3_t$are not available due to the initial condition.

4. CONCLUSION

In this paper we address the dynamic multi-period, multi-modal transportation problem, in whichwe determine at the beginning of every period the shipping quantity of single product via multiple routes by considering the on-hand inventory at the destination, in-transit inventory, and demand just realized at the end of the previous period at the destination, so as to minimize the total cost and maintain the required service level at the destination.

To simplify, without loss of generality, we described the problem using only two routes; namely, fast route and slow route, with lead time of one and three weeks, respectively. We proposed a heuristic algorithm that can determine the optimal shipping quantity by assigning a certain portion(λ=0.8) of the average demand to the slow route. By doing so, we can at least determine the shipping quantity for slow and fast route, while attaining the desired service level.

An attempt has been made to further identify the relationship between optimal shipping quantity andλ, other parameters, which has to be continued in the future study. Generalization of this problem will be to determine the optimal allocation parameter set of λ_i's for i-th route. Many other practical considerations can be taken into account as the future study, such as supply capacity of the product from the origin; or limited transportation capacity of routes, or transportation cost not proportional to the shipping quantity due to discrete nature of capacity of the carrier such as containers.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through theNational Research Foundation of Korea(NRF) funded by the Ministry ofEducation, Science and Technology(grant number 2010-0010730).

5. REFERENCES

Barnes-Schuster, D., Bassok, Y., and Anupindi, R., "Coordination and Flexibility in Supply Contracts with Options", Working Paper, Graduate School of Business, University of Chicago, Chicago, Illinois, 1999.

Bensoussan, A., Crouhy, M., and Proth, J. M., "Mathematical Theory of Production Planning", North-Holland, New York, NY, 1983.

Chang, T. S., "Best routes selection in international intermodal networks", *Computers & Operations Research*, 35, 2877-2891, 2008.

Donohue, K. L., "Efficient Supply Contracts for Fashion Goods with Forecast Updating and two Production Modes", Working Paper, Wharton School, University of Pennsylvania, Pennsylvania, 1998.

Eppen, G. D., and Iyer, A. V., "Improved Fashion Buying with Bayesian Updates", *Operations Research*, Vol. 45, 805–819, 1997.

Gallego, G., and O¨ Zer, O¨ ., "Integrating Replenishment Decisions with Advance Demand Information", Working Paper, Department of Industrial Engineering and Operations Research, Columbia University, New York, NY, 1999.

Hausmann, W. H., "Sequential Decision Problems: A Model to Exploit Existing Forecast", *Management Science*, Vol. 16, B93–B111, 1969.

Hausmann, W. H., Lee, H. L., and Zhang, V. L., "Optimal Ordering for an Inventory System with Dual Lead Times", Working Paper, Department of Industrial Engineering and Engineering Management, Stanford University, Stanford, California, 1993.

Heath, D., and Jackson, P., "Modeling the Evolution of Demand Forecast with Application to Safety Stock Analysis in Production Distribution Systems", *IIE Transactions*, Vol. 26, 17–30, 1994.

Johnson, O., and Thompson, H., "Optimality of Myopic Inventory Policies for Certain Dependent Processes", *Management Science*, Vol. 21, 1303–1307, 1975.

Kim, I. D., "A Heuristic Algorithm for Multi-Period, Multi-Modal Transportation Problem", unpublished Master's thesis (in Korean), Ajou University, 2009.

Lovejoy, W. S., "Myopic Policies for Some Inventory Models with Uncertain Demand Distributions", *Management Science*, Vol. 36, 724–738, 1990.

Scheller-Wolf, A., and Tayur, S., "A Markouian Dual-Source Production-Inventory Model with Order Bands", GSIA Working Paper No. 1998–E200, Carnegie Mellon University, Pittsburgh, Pennsylvania, 1998.

Sethi, S., and Sorger, G., "A Theory of Rolling Horizon Decision Making", *Annals of Operations Research*, Vol. 29, 387–416, 1991.

Sethi, S. P., Yan, H., and Zhang, H., "Peeling Layers of an Onion: A Periodic Review Inventory Model with Multiple Delivery Modes and Forecast Updates", *Journal of Optimization Theory and Applications*, Vol. 108, 253–281, 2001.

Sethi, S. P., Yan, H., and Zhang, H., "Inventory Models with Fixed Costs, Multiple Delivery Modes, and Forecast Updates", *Operations Research*, Vol. 51, 321–328, 2003.

Song, Sang Hwa, "Multi-Period Integrated Inventory and Distribution Planning with Dynamic Distribution Center Assignment", *Industrial engineering & management systems : an international journal,* 5(2), 132-141, 2006.

Yan, H., Liu, K., and Hsu, A., "Order Quantity in Dual Supply Model with Updating Forecast", Working Paper, Department of Systems Engineering and Engineering Management, Chinese University of Hong Kong, Hong Kong, 1998.

Zhang, V. L., "Ordering Policies for an Inventory System with Three Supply Modes", *Naval Research Logistics*, Vol. 43(5), 691–708, 1996.

OPTIMIZATION OF INTER-MODAL INTERNATIONAL LOGISTICS

-AN EXTENDED STUDY OF INTER-MODAL TRANSPORTATION FROM EAST ASIA TO U.S.-

Katsuhisa Okita[1] and Takashi Amemiya[2]

[1]United Parcel Service Company,
3-3-6, Ishida, Minato-ku, Osaka, 545-0006 JAPAN
Email: k-okita2010@ares.eonet.ne.jp

[2]Department of Business Administration and Information,
Setsunan University,
17-8, Ikeda-Nakamachi, Neyagawa-shi, Osaka, 725-7508 JAPAN
Email: amemiya@kjo.setsunan.ac.jp

Abstract: In an evolving environment surrounding international logistics, in place of " multimodal" transportation, "inter-modal" transportation supported by containerization has become the main stream today. Behind this progress is the development of infrastructure, made in response to the users' demand of comprehensive rationalization of logistics, by incorporating efforts made by pertinent parties such as operators of ports, airports, container terminals, maritime company, air carrier and cargo-handling. In this new setup, container transportation plays the key role. It is evident the containerization will be of further importance in the promotion of world trade. There is room for improvement in the inter-modal system. For example, transportation by sea from East Asia to Japan and those by air from Japan to the USA is an uncultivated inter-modal selection. For the sustainable development of highly efficient international logistics, comprehensive study on the operational specifics and management techniques needed to attain optimization of inter-modal international logistics is required. Knowing these variables well and finding their optimal combination will lead to the creation of an optimal solution. And this surely is the attribute that makes inter-modal logistics increasingly important and enables it to contribute to the overall rationalization of international logistics.

Keywords: Inter-modal logistics, Containerization and Optimization

1. INTRODUCTION

Remarkable progress is evident in the environment surrounding international logistics. In the international cargo market, since the last half of the 1960s, containerization of cargos has advanced rapidly, along with the expansion of the container vessels' service routes. Today, most cargos, excluding crude oil and bulk cargo, are carried by linking of sea transport and land transport, and of sea transport and air transport, both using containerized transportation as their medium. This is called "Inter-modal International Logistics." In place of "multi-modal" transportation consisting of more than one mode of transportation and it requires transshipment at respective transfer points, "inter-modal" transportation supported by containerization has become the mainstream concept. This advance has the following as its backdrop: user demands for comprehensive rationalization of logistics which offers consistent and across-the-board control from production to consumption; and in response to this, the development of infrastructure, achieved by incorporating efforts made by pertinent parties. They are operators of ports, airports, container terminals, sea-freighting, air-freighting and cargo-handling. In this development of new infrastructure, container transportation plays the key role. In this study, inter-modal transportation strategy comprising a number of variables is examined from the viewpoints of the providers and the users of inter-modal logistics services. This report defines the existing issues to deliver the possible suggestions through the study.

2. BACKGROUND IN EVOLVING INTER-MODAL INT'L LOGISTICS

Today, most cargos, excluding crude oil and bulk cargo, are carried by linking of sea transport and land transport, and of sea transport and air transport, both using containerized transportation as their medium. This is called "Inter-modal International Logistics." Inter-modal transportation is the freighting system by which cargos are carried from a place of departure to a destination via a combination of sea transportation and rail or road transportation, or a combination of sea and air transportation. In the background of the development of inter-modal transportation realized by liners and airline companies, there was an imperative for the logistics companies to add an alternative to the conventional uni-modal transportation method which enable them to better respond to the higher level of needs of consigners that came about with the changes in environment surrounding them. The following is a list of basic reasons behind the consignors' higher level of needs.

(1) Along with the progress of globalization of manufactures, their investment in countries where inexpensive labor is available increased. As a result of that, consignees started to strongly request for a global transport network while showing their dissatisfaction with the transportation system designed by linking multiple liner shipping routes.
(2) Placing great importance in Just-In-Time system as means to reduce inventory, consignees started to demand quantitative improvement in transport services. In other words, the consignees are strongly demanding the ocean liner companies to provide the more accurate and more regular-scheduled transport services.
(3) Consignees are pursuing improvement of value added to their products brought about by qualitative improvement of transport services.
(4) Consignees are demanding qualitative
(5) Improvement of transport services realized by international collaboration.
(6) As the consignees needs call for the transport services beyond the scope of sea transportation, it is becoming increasingly difficult to come up with the best transporting method using the services available within the framework of the marine conference only. The above trend suggests that movement of mass and regular cargos between specific countries and regions is necessary for the development of inter-modal transportation system. In order to realized the development, good team work between transport service providers to ensure reliable transmission of cargo data and collaboratively structuring of services will be a prerequisite

3. THE INEVITABLE OF INT'L INTER-MODAL TRANSPORATION IN EAST ASIA (for cargos departing from China)

Against the backdrop of recent economic development in East Asia, development of transportation infrastructure such as expressways, ports and harbors, and airports, in the region is rapidly progressing, and the container ships and aircrafts moving within and from the region are becoming larger. This affected the consignees' demands also. They are now demanding the through rate and the comprehensive transport responsibility which covers from the point of receipt of the cargo by the transportation service provider through the point of the arrival at the destination. In response to this, the logistics service providers started to functionally and efficiently combine multiple transportation modes, which has been adding importance to international multimodal transportation system. As well as the leading manufacturers in the world have been shifting their production base from the Southeast Asian region to China, in pursuit of inexpensive labor, European and American retailers, who import a massive volume of commodities via sea freighting, are expanding their scale of procurement from China. These have led to double-digit increase in the number of containers used for sea transportation from China to Europe and the United States every year lately. There has also been a change in export items from China as well, from the textile products to electronics.

(1) Current Situation of Sea and Air Transportation from China to the U.S.A. via Japan
Direct investment by Japanese companies in China has become active since around 1990, and proactively prepared itself to attract foreign investments. In addition to its four free trade zones such as Shenzhen and Xiamen, it set up "open coastal cities," with 14 cities along the coastline. With this, China further developed its infrastructure for attracting foreign investment. In the beginning, Japan observed these Chinese initiatives with prudence, but finally Japanese companies started to move. Of the above 4 free trade zones and 14 open coastal cities, however, Japanese companies most often made their way to one of the three regions encompassing Shanghai, Shenzhen or Dalian. Behind this concentration is the primary purpose of Japanese companies' moving into China: the formation of an "export/assembly base" in an infrastructure that offers "cheap and plentiful labor," and the utmost requirement for such bases is their "port capacity." On the very long Chinese coastline, today Hong Kong has the best port capacity and this has made neighboring Shenzhen and its periphery in the southern part of Guangdong Province attractive locations. Besides Hong Kong, the ports of Shanghai and Dalian are recognized to be first-rate, and these three areas and their peripheries have become the focal points for Japanese companies' investment in China. This tendency to prefer a location with the advantage of great port capacity was shared with companies of other origins, too. They also had the initial intention of establishing an "export/assembly base," while having access to "cheap and plentiful labor," and found the areas around Shanghai, Shenzhen and Dalian most suitable. As a result, a massive foreign investment was concentrated in these areas, which made them grow by leaps and bounds, and having a great impact on China's policy of openness and reform. As China's economic growth in recent years has been remarkable, especially eastern China, which includes Shanghai and its hinterland such as Jiangsu Province and Zhejiang Province, has been receiving the world's intense interest. This region forms a great economic area, with Shanghai at its core. Factories are being built one after another in Shanghai, Suzhou, Wuxi, Hangzhou and Ningbo. Some factories have started operation and product shipment. These cargos are exported from the port of Shanghai and the Pudong International

Airport, bringing the port and the airport an unprecedented booming economy. In terms of air cargo, however, it is not always found to match the ideal logistics envisioned by consignees. This is largely due to the shortage of facility capacity at the Pudong International Airport. In the eleventh year since the opening of the airport, airport facilities cannot fully accommodate the ever-increasing export cargo, resulting in a backlog. Some goods end up waiting for a few days through several days before being moved to the air craft. This situation of export cargo's exceeding the loadage of aircrafts has brought about hikes in air tariffs. Turning to the port of Shanghai, the sea-way entrance, it is evidently one of the busiest ports, with its container transaction volume ranking third in the world, following Hong Kong and Singapore. Shanghai is included in the routes operated by most of the liner route shipping houses. From Shanghai, the inter-modal transport is carried out. Using the bellies of these ships, cargos are transported by sea from Shanghai to nearby ports in East Asia such as Kaohsiung, Pusan, Fukuoka, Osaka, and Tokyo, and then transported by air to the U.S. markets after being transferred to the nearest airport.

(2) Current situation of Japan ports to Sea & Air Transportation
On the one hand, due to the hollowing-out of industry in Osaka, the demand for export air cargo to the U.S. from the Kansai International Airport (KIX) is sluggish. Even compared to the demand level in the airports in China and other countries in East Asia, it is low. On the other hand, it is anticipated that the Kansai International Airport, along with the ports nearby, will come into the limelight again as the transit point for the inter-modal transport from China to the U.S.A., complemented by well-developed ports and harbors with a long history of handling sea-freighted import cargos at the landing ports of Kobe and Osaka since pre-war times. In a given situation of airport/harbor infrastructure and the export cargo market, a need of Sea and Air transportation to U.S. from China via Japan has been raised. In terms of cost and number of days required for transit, for instance, when they are compared between that of Sea and Air transportation and that of sea transportation or air transportation, the cost of Sea and Air is about one-third of air and the number of days required for transit for Sea and Air is about a half of what is required for sea transportation. Due to a rise in airfare after the 1973 oil crisis, the consignees became keen on reducing transportation cost, which resulted in the Sea and Air transportation attracting even more attention than before. The attraction is that, with Sea and Air transportation, mass shipment at low cost is possible and adjusting the timing to supply goods to the market is easy.

4. OUTLINE OF SEA & AIR TRANSPORT (EXISTING ISSUES)

The sequence of sea and air transport starts with sea transport from the agreed place of shipment to the transit point, followed by air transport from the transit point to the destination. Though its history is still short, it stemmed from the arrival of container ships at sea and jumbo jets (B747-F) in the air, and it developed by capitalizing on the strengths of "inter-modal international logistics," such as the low cost of sea transport and the swiftness of air transport. The combined Sea and Air transport has its origin in the field of air cargo transport. It is said that Flying Tiger Line, Inc., a U.S. airline company, started this in the last half of the 1960s, when they air-freighted cargo, which arrived in the west coast of the U.S. after being sea-freighted from Japan, inland and to the east coast of the U.S..
The characteristics of this mode of transport include the following:

1. A smaller number of days required for transport compared to the sea-only mode.
2. Less expensive freight costs compared to air-only mode.
3. Possibility of timely transport of the proper amount in response to commodity demand.
4. Possibility of logistic cost saving, in inventory adjustment, storage fee, etc.

The standard number of days required for Sea and Air transport, though it varies depending on the routes and means of transport, is 7~10 days from Shanghai (by sea) via Japan to (by air) the U.S.A. This means a saving by more than 50% compared to the number of days required for the sea-only mode. The cost is also less – about one-third of air-only transport. The number of days required for transport has a great impact on determination of the freight cost. The shorter the time is, the more expensive, naturally. As the Sea and Air transport gives flexibility to transport by getting the best of both sea transport and air transport and by being able to respond to the specific nature of the cargo, its condition, and the freight costs, it is called the "third logistics" in the field of international logistics and its potential as the intermediate mode of transport between the Sea and Air modes is increasingly recognized. While the primary cargo for Sea and Air transport from east China, for example, includes electric appliances, electric components, office appliances, audio equipment and semiconductors, it is anticipated that transport not only from this region but also to this region will increase. It is apparent that many transnational corporations from Japan, Europe and the U.S., as well as other nations, will continue to set up operations overseas and local production will continue to grow. This leads to the expectation of increased transport to China to support local production there with supplies of parts.

5. COMPLEMENTARY FUNCTION AND SUBSTITUTE FUNCTION – SWITCHING OF MODES AS A PROACTIVE STRATEGY

In the sense of innovation in transportation, the container-sea transportation service can be said to be in the same dimension with the air transportation service. Because they are greatly different from each other in terms of speed, it should be most appropriate to regard them as being complementary rather than as being competitive to each other. Hypothesizing the fundamental complementary and partially substituting relationship between the container-sea transportation service and air transportation service is meaningful as the resultant hypothesis can be used to explain the real situation logically. However, there are opposite views to this. Some people see that the sea transportation service and air transportation service are generally substitutive to each other. It is said that the sea transportation is on a losing streak against air transportation. This is due to the fact that air transportation service is compared to sea transportation service in general, not to the container-sea transportation service. This is what is done in the discussion focusing on the change of the ratio of "switching from sea to air cargo". The ratio of "switching from sea to air cargo" identifies the ratio of the value traded by air freight in the total value of trade of our country, and it is indicated in percentage. For example, the ratio of "switching from sea to air cargo" for Japan's export, which indicates the ratio of the value of air freighted exports in the total value of export from Japan, has consistently increased in the past. In contrast, the trend has been consistently downward for the share of sea freighting (or one might call it the ratio of "switching from air to sea cargo" for Japan's export), which is obtained by subtracting the ratio of "switching from sea to air cargo" for export from 100%. As described above, the ratio of "switching from sea to air cargo" and the share of sea transportation are, by definition, constantly in an inverse correlation, and furthermore, the ratio of "switching from sea to air cargo" is on the increase. This leads to an interpretation that the sea transportation service and air transportation service are substitutional services, and that the air transportation service is a normal service and the sea transportation service is an inferior service. This understanding is not wrong in so far as the foregoing definition is applied, but as far as the logistics innovation is concerned, the sea transportation service that is to be compared to the air transportation service must be limited to the container-sea transportation service. As explained above, the Sea and Air service, a mode of international multimodal transportation, is of air and container-sea transportation, and this will continue to grow as the air and container-sea transportation services will supplement weaknesses of each other, and as this mode of transportation will meet the intermediate values of air and container-sea transportation services in terms of the number of transit days, freight tariff, and weight and volume of exported cargos.

6. STRATEGICAL BEHAVIOR OF SERVICE PROVIDER

There are several parties involved in the international logistics. The freight forwarders link the consignees and the transportation companies by carrying the shipment door to door. While the carriers (ocean shipping companies and airline companies) move the cargos along the transportation routes, the forwarders are active, doing cargo handling for example, at nodal points of each link. One of the activities of the international logistics service providers involved in air freight is the logistics strategies development performance by the forwarders. Unlike sea freights, most air freights are booked by forwarders due to historical and institutional reasons. As a result, the air freight forwarders' logistics-strategies-development performance will be the center of focus for the consigners. Competition in product development among these freight forwarders caused diversity in transportation routes and the number of days required for transit, which has led to progress in freight cost reduction and further to establishment (as a conventional product) and growth of Sea and Air transportation. In the perspectives of the users, being able to choose the transportation method that can minimize loss/damage in relation to the following will be important – freight cost premium when the consigners decide to switch from the slow transportation mode to the fast transportation mode; or an adverse impact on the future businesses or penalty accompanied by a breach of contract caused by a delivery delay. One of the causes of increased demand in the combination of air and Sea and Air transportation is, for example, that the consigners are choosing the conventional air transportation paired with the Sea and Air transportation to timely replenish the consignees' inventory rather than risking back order resulting from a drastic decrease in their inventory. To be better prepared for the further increase in demand of international multimodal logistics in future, the following improvements are desired.

6.1 Development of Ports and Harbor/Airport Infrastructure

As the cargos landed at each port are reloaded on trucks to be transferred to the airport, due to the current situation in Japan where there are a number of traffic problems remaining to be solved, there is the possibility of being caught in traffic congestion, leading to a delay in reaching the airport. The road infrastructure, around the nodal points between the ports and airports in particular in this case, must be developed.

6.2 Round-the-Clock Operation at Ports and Harbors

It is said that ports in Japan are small in general. Other major ports in East Asia, such as Hong Kong, Kaohsiung, and Pusan, are even smaller against the total cargo volume to export and import. Due to this, they operate with 24-hour cargo-handling to effectively use the terminals. A similar 24-hour operation is necessary and expanded to Japanese ports to increase their efficiency.

7. Promotion of Effective Logistics

The recent developments also include a new movement to integrate production and logistics. Though production and logistics once took separate courses of development, the new approach is to integrate these functions while maintaining the optimal relationship between them, to realize customer satisfaction and low-cost operation. When it comes to world trade and international logistics, there are so many parties involved. Some of them are business operators and some are of the public sector. To make it more complicated, there are different regulations, customs and business practices in different countries and regions. All of these make the requirements of paperwork more than just complex, and the burden and the cost of processing are huge. To realize ideal international logistics, at least in terms of processing import/export and other needed documentation, a system of total optimization encompassing both public and private sectors must be built. It is anticipated that, on the one hand, stronger demands from the customers for rationalization in logistic activities will be made, and on the other hand a demand for the timely and accurate provision of information, as well as proposals for optimal solutions, will be made. For the former, containerization is expected to continue to play an important role.

8. CONCLUSION

As the long-term policies generally made by advanced nations are designed to reduce imports of fuel and raw materials, and instead increase those of finished products, it is evident that containerization will be of further importance in the promotion of world trade. For the sustainable development of highly efficient international logistics, further study on the operational specifics and management techniques needed to attain optimization of inter-modal international logistics is required. Examination of a number of variables and their combinations, which constitute inter-modal logistics as seen from the viewpoints of both the service providers and the users in this study, suggests ways to improve the throughput, cost effectiveness and other aspects of logistics, and thus the level of customer satisfaction. Knowing these variables well and finding their optimal combination will lead to the creation of an optimal solution. And this surely is the attribute that makes inter-modal logistics increasingly important and enables it to contribute to the overall rationalization of international logistics.

9. REFERENCE

[1] Amemiya, T., Takeyasu. K., Masuda, S.(2006). New Mathematics for Management & Economics (in Japanese) Chuo-Keizai-sha, Tokyo, Japan
[2] Suzuki, A. (2001). Theory and Practice of International Physical Distribution (in Japanese). Seizando Shoten, Tokyo, Japan
[3] Sakakibara, S. (1999). Economy of Air Transportation (in Japanese). Koyo Shobo, Tokyo, Japan
[4] Kinoshita, T. (1999). Theory and Practice of International Air Cargo Transportation (in Japanese). Dobunkan, Tokyo Japan
[5] Sakakibara, S. et al (1999). Inter-modalism (in Japanese). Keiso Shobo, Tokyo, Japan
[6] Saito, T. (1997). Structure of Transportation Market Policy (in Japanese) Chuo-Keizai-sha, Tokyo, Japan
[7] Maeda, Y. (1998). Overview of Transportation Economy (in Japanese). Koyo Shobo, Tokyo, Japan
[8] Miyashita, K. (1999): Marine Transportation (in Japanese). Koyo Shobo, Tokyo, Japan
[9] Masui, K. (1999). Theory of Air Transportation (in Japanese). Koyo Shobo, Tokyo, Japan
[10] Umezu, K. et al (1999). New Trends of World Trade (in Japanese). Koyo Shobo, Tokyo, Japan
[11] Yamagishi, H.(2004). Theory of Sea Container Logistics (in Japanese). Seizando Shoten, Tokyo, Japan

THE FUTURE OF SEAPORT HINTERLAND NETWORKS

Albert Veenstra* and Rob Zuidwijk

Department of Information and Decision Sciences
Rotterdam School of Management
PO Box 1738
3000 DR Rotterdam
{aveenstra, rzuidwijk}@rsm.nl}

Abstract: In this paper, a vision is developed on the future of seaport hinterland networks. The innovative concept of the Extended Gate is discussed in detail, based on the DryPort concept which has already received considerable attention in practice and in the scientific literature. We then identify a number of business challenges related to the implementation of the Extended Gate concept. Based on a literature survey, we identify research challenges that we aim to address in future research projects, such as the Dinalog project ULTIMATE. We also assess the economic impact of the Extended Gate concept once implemented. This chapter is dedicated to Jo van Nunen, who cherished the port of Rotterdam as a laboratory for innovation and scientific research, and who has played a pivotal role in the application of scientific research in logistics and supply chain management.

1. INTRODUCTION

The Netherlands has been a trading nation for centuries. As a result, the country has a strong fiscal-financial infrastructure, a strong and innovative logistics industry, and, last but not least, one of the largest ports in the world – the Port of Rotterdam. It is therefore not surprising that the Netherlands plays an important role in a large number of global supply chains. Supply chain innovations often find their focal point in the Netherlands. This makes the Netherlands, and the Port of Rotterdam in particular, an excellent laboratory of logistics and transport research. Jo van Nunen, to which this chapter is dedicated, has always strongly supported the idea of the port as a laboratory where innovative practitioners and scientists could work together to develop new concepts. This chapter discusses a recent innovative concept in the port that will carry our ErasmusUniversity tradition of strong academic, applied research in the port well into the future.

The port community in Rotterdam is continuously facing challenges in moving cargo through the port: congestion, pollution, delays due to administrative burdens, instability of hinterland connections, fierce – if not cutthroat – competition between transport operators and terminals, and a search for the relationship between added value activities and the cargo handling process.

In the years 2004 and 2005, when the unpredicted surge of containers from China reached Europe, the various container terminals became so congested that there was a strong need for a breakthrough innovation. As a response, the idea emerged that a seaport terminal should be able to push blocks of containers inland pro-actively, to alleviate congestion. The Rotterdam deep sea terminal operator ECT, assisted by various master students and researchers from Hogeschool Rotterdam and ErasmusUniversity, developed this idea into the concept of the Extended Gate. In this concept, the terminal gate is extended to include selected hinterland locations, which allows for the movement of containers to those inland locations without prior involvement of the shipping company, the shipper, and the receiver. From 2007 onwards, this concept became operational on the link between the ECT terminals in the Port of Rotterdam and the inland terminal TCT Venlo. ECT is now, in collaboration with ErasmusUniversity and Eindhoven University of Technology, working on the implementation of this concept into a network of inland locations.

This chapter concerns the research challenges that we face in developing the extended gate concept into a network concept. This work will be part of the project ULTIMATE (Efficient Multimodal Hinterland Networks – new concepts for design and operations) supported by the Dutch institute Dinalog. Jo van Nunen was one of the founding fathers of this national institute for logistics and supply chain management. This chapter addresses the research agenda for this project, and may serve as a vision paper for hinterland network structures in Europe for the next decade.

2. THE DRYPORT AND THE EXTENDED GATE CONCEPTS

In sea port hinterland intermodal[1] transport networks, as depicted in Figure 1.1, different organizations, such as terminal operators, freight forwarders, information service providers, infrastructure managers, shippers, and receivers, all aim to contribute to a better performance of the overall supply chain. Terminal operators, for instance, are more and more involved in linking sea terminals with inland terminals, or linking terminals with final destinations in the supply chain, such as warehouses, and henceforth they shift their activities from being a "node operator" to a "flow operator". It enables them to better connect with shippers and receivers in the network. In this vein, ECT has developed a network of inland terminals in Venlo (NL), Duisburg (D) and Willebroek (B). ECT has been offering rail services between Rotterdam and Venlo for a number of years, and this service turns out to be reliable, fast, and cheap; trucking is only marginally faster, but more prone to congestion on the motorways, and more expensive. A major driver for this success is the high frequency of three or four daily departures. Another major driver is the administrative integration of the service across the deep sea terminal – transport – inland terminal link. Furthermore, four independent terminals in the Brabant region have set up a joint subsidiary – called Brabant Intermodal - to coordinate shipments from the deep sea terminals, again creating conditions for larger shipments with higher frequencies.

Shippers increasingly demand efficient and sustainable intermodal transport services. One example is Proctor and Gamble, who have developed an environmentally friendly washing gel that can wash at 15 °C. To move this project into their sales channels in Europe, they also demand environmentally friendly (read: intermodal) transport solutions from the seaport to their various warehouse locations across the UK and the European continent. So the initiatives by ECT and Brabant Intermodal are exactly what companies such as Proctor & Gamble are looking for.

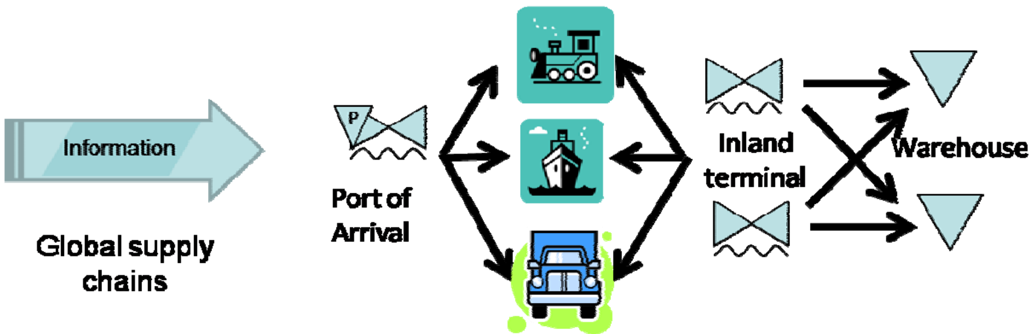


Figure 1.1. Sea port hinterland intermodal transport network

We now elaborate on the so-called Extended Gate concept in seaport hinterland intermodal networks. Below, we will develop the concept of the Extended Gate from the more basic concept of the DryPort.

Next to the ongoing issue of trade growth and congestion in terminals, the development of newly industrialised countries, and their need to gain access to the global transport system, has been instrumental in the development of the so-called DryPort concept. This concept is based on the idea that not all industrial and economic activities have to take place close to seaports (as is common practice in many developing countries), but that good infrastructure and inland nodes can help accommodate trade growth, and bring regional development inland. As a result, UNESCAP has adopted the DryPort development as one of their main strategic objectives; see www.unescap.org, and the work of UNESCAP economics affairs officer Wang, Tengfei.

Roso et al. (2009) introduce the DryPort as a driver for alternative transport network configurations. They observe that Dry Ports serve to bundle cargo and mostly serve a community of cargo interests and operators. The main purpose of the DryPort is to transfer activities from the seaport to the DryPort to relieve congestion, and achieve other benefits. Roso et al. (2009) specifically state the potential for modal shift as one of those benefits.

The view on Dry Ports from a seaport economics point is put forward by Notteboom and Rodrigue (2005), who see the development of hinterland networks as a new dimension for competition between seaports. Much of the literature in this area, see for example (McCalla, 1999), focuses on the reasons why hinterland networks emerge, and on the relationships between the networks of various seaports that are in competition.

In the meantime, in another strain of work, much attention has been given by practitioners in seaports and related areas to develop hinterland networks, often centred around a developed, or developing Dry Port. For the European

[1]For ease of exposition, we refer to intermodal transport as the transport of containers using several modes of transport. However, one may reserve this term for the consecutive use of different modes along a transport lane, and refer to the use of several modes of transport in parallel on a single transport leg as multimodal transport.

Commission (2001, p. 59), a Dry Portis "an inland terminal which is directly linked to a maritime port". Leveque and Roso (2002) provide an alternative definition that is a bit more precise: "A Dry Port is an inland intermodal terminal directly connected to seaport(s) with high capacity transport mean(s), where customers can leave/pick up their standardised units as if directly to a seaport" (our emphasis). The definition indicates the purpose of the Dry Port: it acts as an extension of the seaport, both for export and for import traffic.

Numerous examples exist of Dry Port-like arrangements. Roso et al (2009) mention a number of examples, and the terms under which they were presented: Inland container depots (India), Güterverkehrszentra (Germany), Enhanced remote transit sheds (UK). More specifically, examples are (non-exhaustive):

1. TCT Venlo, the Netherlands with ECT container terminal in the Port of Rotterdam
2. Virginia Inland Port with Hampton Roads seaport
3. Alameda corridor between LA and Long Beach, and its connections with intermodal terminals
4. Enfield, close to Sydney Botany Bay seaport
5. Isaka Dry Port and Dar Es Salaam in Tanzania
6. Caslada, Madrid, with seaports in Valencia, Bilbao, Barcelona and Algeciras
7. Toulouse Logistics Activity Zone, and the Port of Barcelona

In Europe, the Port of Rotterdam, the Port of Antwerp, ports in the UK, and ports in Germany and Spain have all developed the Dry Port concept to some degree. In the UK, the Customs arrangements already facilitate completely paperless procedures in ports, including the pre-release of containers, but this is based on a feature of the UK Customs system that may disappear with the introduction of the new Community Customs Code. In the other cases, the implementation is mainly within a single Member State. Crossing borders with the Dry Port concept remains a challenge.

We now further develop the Dry Port concept in to the Extended Gate concept. The Extended Gate concept incorporates some of the natural consequences of the Dry Port concept, such as integral network design and direct operational control in the transport network between the sea terminal and the Dry Port. However, as we will see, the Extended Gate distinguishes itself by putting more emphasis on coordination and control of the flows in the intermodal hinterland network.

Even if seaport terminals do not control the containers coming in by sea, they can develop instruments and – in collaboration with other stakeholders of the container-based supply chain – influence the flow of containers to the final destinations.

The Extended Gate concept establishes the point of delivery from the perspective of the shipper/receiver at the appropriate inland terminalor, when possible, at the final destination such as a distribution centre of the shipper/receiver. Essentially, the gate of the sea terminal is now placed at any of the inland terminals in the network (Visser et al., 2007). Cargo interests agree to pick up their containers at the inland terminal location, and the final leg of the journey is arranged according to this agreement by a terminal or another operator. This implies that the shipper/receiver will directly deal with a wide variety of inland terminals rather than with the sea terminal. This delivery at the inland terminal, or even at the distribution centre, is offered as an additional service to the customer. Basically, inland terminals located in economic centres are the most suitable for this purpose, because they could facilitate the flow of import and export containers and facilitate the flow of intercontinental cargo as well. Extended Gate terminals can also build on the function of an intermodal platform, from which containers could be carried on to other, more distant locations, by rail or inland barge. This will help these inland terminals build their service portfolio and develop into true intermodal service providers. In some cases, the customs regime is extended to the inland terminal as well, creating additional value of the service package.

The delivery of the aforementioned intermodal services comes with serious and unexplored business challenges in the areas of collaborative planning and information sharing, intermodal network management in global supply chains, and business model development.

A major challenge comes from the actual usage of newly available information. For instance, if information along the supply chain is integrated and a smaller number of parties involved decide on operational priorities, then the final destination and required timing of delivery of a container (which is typically known by the inland terminal) can be linked to the Estimated-Time-of-Arrival information at the deep sea terminal to pre-arrange containers in the stack differently, and to consolidate containers shipments. Current research has not yet addressed this coordination opportunity, as information is not shared and substantially more parties are involved with coordinating the network from the arrival in the sea port until delivery at the receiver's warehouse.

One of the crucial conditions for the development of efficient hinterland networks in Europe is the availability of the right information on goods that are arriving from overseas to relevant stakeholders. This includes information on the nature of the goods, quality, health and origin certificates, safety and other handling instructions, destination, shipper, receiver, intended mode of hinterland transport, and required arrival date and time. This information is usually in the hands of freight forwarders, and of the owners of the goods. Currently, such information becomes available to container terminals or hinterland transport operators only at the very last moment, while they need it in advance to better plan their operations.

In addition to the availability of information, an important condition for the Extended Gate networks to function is that the inland terminals are equal partners to the seaport terminals in terms of the quality of their logistics performance, the information management capabilities, their internal terminal management systems, their account management, and customer relationship management, security, and customs status. The conditions the Extended Gate network has to comply with have to be derived from the supply chains they facilitate. A thorough understanding of the supply chain requirements for transport network integration is thus required.

In networks based on Extended Gate concepts, specific operational and analytical problems arise that have to do with the optimisation of operational activities across nodes and links in the network. Examples are the reliability of turnaround times of trains and barges in the port of Rotterdam for the performance of trains and barges in hinterland corridors, the requirement for more flexibility in stacking operations in both sea and inland terminals due to peaks in traffic, and the requirement to show efficiency benefits in operational processes in the sea terminal to offset investments elsewhere in the network.

The current examples of Dry Ports/Extended Gates, except perhaps the example of the Alameda corridor, are about one single connection between a Dry Port and one or more seaports. There are still no examples where one seaport is connected to a network of Dry Ports. There is little experience with issues of integration, competition and collaboration within such networks.

3.INTERMODAL TRANSPORT IN GLOBAL SUPPLY CHAINS

Global supply chains are becoming increasingly complex; see for instance (Deloitte, 2003). Companies are relying more and more on global sourcing of semi-finished and finished goods, and are confronted with progressively more complex demands of customers in terms of reliability and timeliness of deliveries. In addition to this, there is a strong recent interest in supply chains that are both efficient, reliable and sustainable. Bowersox et al. (2005) list ten supply chain mega-trends, which suggest a transition towards collaborative business networks as flexible, dynamic, customer oriented networks in which economics, quality and sustainability are evenly matched, and in which information is managed to establish online planning, virtual communication and demand oriented processes (Vervest et al., 2009).

The supply chain management literature recognises the transport function as an element in the integration of supply chains; see for instance (Morash and Clinton, 1997). However, the development of integrative activities at the transport level receive very little attention. Stank and Goldsby (2000) discuss transportation management within the context of existing transport possibilities, while Mason et al. (2003) discuss the integration between transportation and warehousing. Rodrigues et al. (2008) present a study on taking into account transport uncertainty in supply chains. Their analysis points at a considerable sensitivity at the supply chain level due to uncertainties in the transport and logistics function, and also to the difficulties in identifying the exact source of these uncertainties. This is in line with similar statements of Rodrigues (1999), who, almost a decade earlier, observes this same phenomenon. He argues that synchronization of (container) terminal activities across supply chain networks is the main source of efficiency gains. In addition, Bruinsma et al. (2000) show that the quality of transport networks, especially the potential for intermodality, can attract new business activities.

The merits of finding a better integration of transport and logistics in the supply chain seem evident. On the integration of modal transport networks, Van Geenhuizen (2000) argues that the concept of transport network interconnectivity has not been defined very well. She also shows that a demand or user perspective is required in network interconnectivity analysis, to maintain a sufficient relevance of the interconnected networks for users. Again, from this perspective of intermodal transport economics, Roson and Soriani (2000) state that the terminals incur the largest changes.

The research field on intermodal transport networks takes concepts from network theory and applies these to the networks that emerge when seaports are connected to inland terminals or when inland terminals are interconnected. Bontekoning et al. (2004) provide a literature overview on intermodal networks, arguing for the definition of intermodal networking as a separate research field. Their work, as much of the work in this area, is restrictive because it disregards inland shipping as a mode of transport. Another example of the work in this area is Janic (2007), who develops a costing model for intermodal and road freight transport networks. The comparison between the two configurations of the model (intermodal, and road freight only), provides him with a framework for European transport policy analysis.

In addition to transport networks, there is an extensive body of literature on intermodal terminal operations. See for a recent survey Stahlbock and Voss (2008), Much of this work, while relevant and scientifically challenging, focuses on internal processes at the terminal. If connections with the logistics chain are considered, it is often an allocation process that takes place at 'the edge' of the terminal that is studied. An example is the relationship between a terminal and landside transportation. There are some exceptions, such as Veenstra and Lang (2009), who study the interaction between ship arrivals and terminal operations, and in particular the impact of capacity shortage of the terminal on the proliferation of delays of ships.

4. GAP ANALYSIS: RESEARCH CHALLENGES

There is a need to create a knowledge base in support of the development of intermodal networks that connect sea- and dry ports, and customer locations inland. The knowledge base should address a variety of issues that will arise in the design, planning, and operation of these networks, in line with the business challenges described in Section 1.2 and the available literature as discussed in Section 1.3.

1. Intermodal networks that incorporate a sea port and multiple dry ports have some specific characteristics that should to be recognized and exploited in their design. Moreover, the existing literature does not address by what mechanisms such inland networks are developed. The network design problem, and the role of major drivers and determinants such as supply chain requirements, and the use of information have not been studied in this context.
2. Quantitative models developed so far that support the design of intermodal networks and services on the network are focused on design principles that are physical in nature, such as transport infrastructures, transport mode characteristics, frequency of service, etc. Apparently, the information infrastructures requirements are considered to be a derivative of network service design and are assumed to be fulfilled. There is a need to further understand how the extent to which information infrastructures are available impacts the design of intermodal services on networks; see for example (Zuidwijk and Veenstra, 2010).
3. The performance of the network may benefit from the integration of the planning of operations at the seaport and dry ports, and the planning of intermodal services on the network. There is a need for methods that quantify the benefits of integrating these planning processes. However, the integrated planning in transport networks across transport modes, and the embedding of these partially integrated networks (read: Extended Gates) into supply chains is an area that remains largely unexplored.
4. The design, planning, and operation of an intermodal network have considerable impacts on the societal and natural environment. On the other hand, stakeholders such as port authorities and governments may impose e.g. environmental requirements that need to be complied with, and in the design, potential regulations need to be anticipated. The modal split requirement for container flows from Maasvlakte 2 is an example. There is a need to incorporate such stakeholder requirements in the modeling that aims to support decision making in the intermodal network.
5. The logistics and supply chain literature deals extensively with the managerial consequences of integration, the legal consequences of integration of (transport) activities and the shifting roles of business parties have only been studied in isolation of logistics and transport practice. Legally very complex situations can occur when a terminal operator also engages in transportation, and in freight forwarding activities, which is the case in Extended Gate concepts. In the intermodal transport literature, topics such as the legal framework for international intermodal transport are in urgent need of further research; see the recent PhD thesis by Hoeks (2009).
6. Most of the Dry Ports use rail as the alternative transport mode. Inland shipping hardly plays a role. In addition, in many cases, the aim of Dry Port concepts is to (only) replace road cargo with rail cargo. Tri-modality does not play a role in the existing examples, nor in the academic literature, but is an important core competence to achieve all potential benefits of Extended Gate concepts. Furthermore, the existing literature takes an "either/or" approach when considering alternative modes, while the joint use of multiple modes is paramount to the success of these networks.
7. Finally, the business case for Extended Gate-type concepts, and possible business models need further attention. The current scientific contributions on the interrelation of transport markets for various modes, on pricing models for multimodal services and on the definition of and demand for network coordination services are quite marginal. Another complicating factor is that in ports, the governance of innovation and the solution of collective action problems rests with a separate entity called the port authority. Thus, the role of port authorities in the development of new multimodal hinterland services, that influence the competitive position of the port in the long run, also deserved attention. The methods that support the design of the intermodal network services should incorporate the business models through which intermodal services create value. For example, demand for transport should not be considered a given but dependent on the pricing of service packages. Moreover, the demand for transport should be recognized as heterogeneous, notwithstanding the standardization of the physical processes of container transport.

5. ECONOMIC IMPACT ASSESSMENT OF THE EXTENDED GATE CONCEPT

We anticipate that the following economic and social benefits can be achieved when the Extended Gate concepts is put in place:

1. The integrated networks with Extended Gates will lead to a more reliable, predictable, and possibly quicker service that facilitates that part of global supply chains where time pressure is the main driver.
2. If the efficient flow of containers into the hinterland is facilitated by Customs, it will also contribute to a better distribution of customs inspection activities to hinterland locations, instead of concentrating those activities in an already congested port area. Visser et al. (2007) also argue that a lot of activities in the sea ports, such as the customs clearance, could be shifted to a hinterland port.
3. Compared to the current situation, container hinterland transport could rely much more on inland waterways and rail services. Given that there is usually enough time to move the container to the inland location, they can be shipped in large quantities by block trains and – if feasible – by barges to inland intermodal terminals. The alternative of road transport can be reserved for the cases under time pressure. Zuidwijk and Veenstra (2010) have developed a model to study the economic, security, and environmental impacts of such solution directions. These solution directions create new demand for more environmental-friendly modes of transport, and contribute to a better spread in the use of existing transport infrastructures.
4. The Extended Gate concept would contribute to decongesting seaport terminals, which allows these terminals to enhance their performance (Woxenius et al., 2004).
5. A decreased use of trucks for container hinterland movements also reduces the road congestion around the seaport areas especially during peak times; see (Roso, 2007).
6. Efficient intermodal hinterland networks will play an important role in location decisions of foreign companies. Making these networks more efficient will therefore increase the relative market share of the Netherlands in the Hamburg-Le Havre range.
7. Inland terminals are required to develop the scope of their operations to satisfy logistics needs of their clients. The supply chains are becoming more complex, so the final customers will need integrated services and capabilities to respond to this increasing complexity (Notteboom and Rodrigue, 2004). This in turn requires more skills and knowledge in the logistics sector, and may create more jobs in the hinterland.
8. Logistic Service Providers/shippers in the hinterland will gain more visibility en can therefore increase the reliability in their supply chains.

We have valued these impacts with some preliminary calculations. The purpose of these calculations is to see how much value added could be generated if the extended gate concept is adopted widely in the port hinterland. The economic impact of the work in the current project consists of the following components:

1. Offering Extended Gate services implies offering high-frequency services to customers of the network upon request, implying a premium service for customers, and hence providing a higher added value.
2. The improved coordination in intermodal transport chains will lower generalised transportation costs for Dutch consumers and companies.
3. The development of intermodal hinterland networks will help shift cargo from road to rail and inland shipping and achieve a better use of existing road infrastructures during non-peak times.
4. The approach to develop seamless hinterland flows of containers will alleviate congestion around terminals in the seaport,
5. The improved intermodal chains will allow the Netherlands to capture a larger market share in North-West Europe's container flows and logistics activities.
6. The development of hinterland nodes as key logistics platforms will have a positive impact on employment creation around logistics centres.

Below, some relevant data and crude estimates of the impact are provided:

1. About 6.5 mln TEU transit Dutch seaports, with an expected growth to 12.5 mln TEU in 2020. Total generalised transport flows: €500 per TEU (Ecorys, 2003), which amount to a total of €3.25 bln. We expect a reduction of 5% on total generalised costs.
2. Modal split: 57% road, 30% barge, and 13% rail[2]. We expect a modal shift of 15% towards rail and barge.
3. Congestion costs on the A15 highway as estimated in 2007: €10 mln per year. Strong growth is expected if no action is taken. This project aims at reduction of congestion costs, but also at improved coordination of trucks which will increase value added per driver.

[2]We acknowledge that different figures are used in different sources; these figures include all container flows, also short sea transport flows.

4. Lock in market share increases of at least 1 percent point for Rotterdam due to efficient hinterland networks. The current market circumstances show market share volatility of up to 5 or more percent points; in 2008, Rotterdam had a container market share of 24%, and in the beginning of 2010, the market share was above 30%! The Extended Gate concept is expected to lock a substantial share of this otherwise dynamic market share. This then becomes a sustainable contribution to the Dutch economy.
5. Employment in Transport and Logistics in the Netherlands was 746.000 persons, with an added value of € 40.2 bln. We assume that employment will grow proportionally with the increase in value added.

Table 1.2. Economic impacts Extended Gate concept

Type of impact	*Value (€ - unless otherwise stated)*
Reduction generalised transport costs	162.5 mln
Modal shift	24.38 mln
Reduce congestion	1.5 mln
Lock in market share (1%)	216.67 mln
Create employment	4000 jobs
Total Impact	**405.05 mln**

6. CONCLUSION: A VISION ON FUTURE INTERMODAL NETWORKS

We finish this paper with the formulation of our vision on future hinterland networks. The analysis of the Extended Gate concept and the envisaged results of the Dinalog project ULTIMATE, as well as the planned redevelopment of the Port Community System in the Port of Rotterdam will have consequences for the structure of hinterland networks in Europe. The main characteristics of our vision are:

1. Hinterland networks will connect seaports selectively with main transport and logistics hubs in the hinterland. The selection of these hubs , and the characteristics and requirements for hubs to qualify for inclusion in hinterland networks is one of the main challenges for the regions in Europe.
2. Seaports will rely heavily on these networks to move cargo into the hinterland, primarily by barge and rail. This means that trucking in and around seaports can be reduced substantially, if there is political will to develop barriers for trucks to enter the port. Pay as you go systems on motorways into the ports for trucks, dedicated truck taxation, minimum driving distance requirements, and other schemes can be adopted for this purpose. See also the Clean Truck Program in the ports of Los Angeles and Long Beach; see www.polb.com andwww.portoflosangeles.org.
3. Seaports will revert back to nodes for cargo handling and transhipment, and will lose their role as logistics nodes. Distribution parks in port areas will disappear, and the available space can be used for other, more directly port related activities.
4. The inspection and supervision regimes of Customs, phytosanitary services, and other inspection agencies will also be moved into the hinterland. This is already the case for many of these inspections, but the administrative processes in the port are not completely optimised with this development in mind. This is an ongoing process.
5. Inland terminals in the hinterland will bear the brunt of the challenges that remain in efficient multimodal hinterland networks: they will sometimes face considerably more serious operational pressure than heretofore, and part of the congestion and other negative effects from the seaports will now materialise around inland terminal locations. It is up to regional authorities to play a role in the mitigation of these effects.

This vision is depicted in Figure 1.2 below. This figure portrays a conveyor belt system between seaports and inland terminals, where the conveyor belt is secured entirely with appropriate technology. Quality standards exist for the selection and performance evaluation of the inland terminal locations.

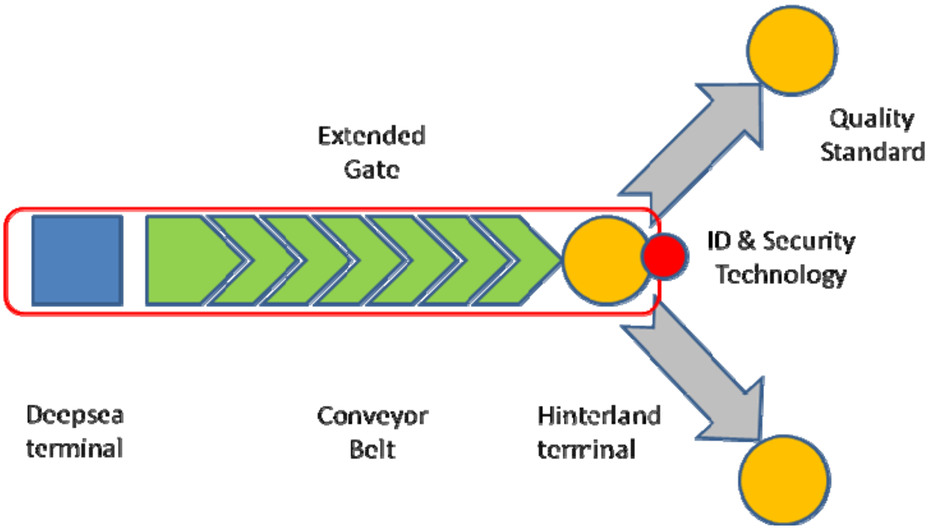


Figure 1.2. Representation of the industrialised extended gate concept
(courtesy of dr R. Mertel, Kombiconsult GMBH, Germany)

If this development will dominate hinterland transport in European seaports, then this sets an agenda for inland terminal managers, transport service providers and policy makers. The quality of performance of inland terminals in Europe will have to be raised to the level that they become equal partners of seaport terminals. In addition, local governance structures will have to be adapted to be able to deal with some of the seaport problems that will emerge in the hinterland. National and European authorities will have to have another look at their supervision arrangements to be able to optimise personnel and resources to keep a check on incoming cargo of all kinds. Additional measures that rely on pre-arrival checks in countries of origin (see for instance the current development vis-à-vis the Import/Export Control System) would need to be implemented and extended. New business partnerships between terminal operators and multimodal transport service providers will have to be established that fit in the current and future legal frameworks for transport in Europe, but that also create new flexibility to offer integrated services in the multimodal hinterland networks.

7.REFERENCES

Bontekoning, Y.M., Macharis, C., Trip, J.J. (2004). Is a new applied transportation research field emerging? – a review of intermodal rail-truck freight transport literature. *Transportation Research* Part A 38: 1-34.

Bowersox, D.J., Closs, D.J., Stank, T.P. (2000). Ten Mega-trends that Will Revolutionize Supply Chain Logistics. *Journal of Business Logistics* 21(2): 1-15.

Bruinsma, F., Gorter, C., Nijkamp, P. (2000). Multimodal Infrastructure, Transport Networks and the International Relocation of Firms. *Transportation Planning and Technology* 23: 259-281.

Dekker, R., Asperen, E. van, Ochtman, G. (2009). Floating stock in FMCG supply chains: using intermodal transport to facilitate advance deployment. *International Journal of Physical Distribution and Logistics Management* 39: 632-648.

Deloitte (2003). Supply Chain Complexity. Deloitte Research Global Manufacturing Study.

European Commission (2001). Freight intermodality. Transport RTD Programme of the 4th Framework Programme, Luxemburg.

Hoeks, M.A.I.H. (2009). Multimodal transport law. PhD thesis, Erasmus School of Law.

Janic, M. (2007). Modeling the full costs of an intermodal and road freight transport network. *Transportation Research* Part D 12: 33-44.

Lang, N., Veenstra, A.W. (2009). A quantitative analysis of container vessel arrival planning strategies. *OR Spectrum* 32(3): 477-499.

Leveque, P., Roso, V. (2002). Dry Port concept for seaport inland access with intermodal solutions. M.Sc. thesis. Department of Logistics and Transportation, Chalmers University of Technology.

Mason, S.J., Rivera, P.M., Farris, J.A., Krik, R.G. (2003). Integrating the warehousing and transportation function of the supply chain. *Transportation research* Part E 39: 141-159.

McCalla, R. (1999). Global change, local pain: intermodal seaport terminals and their service areas. *Journal of Transport Geography* 7: 247-254.

Morash, E.A., Clinton, S.R. (1997). The Role of Transportation Capabilities in International Supply Chain Management. *Transportation Journal*, Spring: 5-17.
Notteboom, T., Rodrigue, J.-P. (2004). Inland Freight Distribution and the Sub-harborization of Port Terminals. In Licheng, S., Notteboom, T. (Eds.), Proceedings of the First International Conference on Logistics Strategy for Ports (pp. 365-382), Dalian Maritime University Press, Dalian (China).
Notteboom, T., Rodrigue, J-P. (2005). Port regionalization: towards a new phase in port development. *Maritime Policy and Management* 32(3): 297-313.
Rodrigue, J-P. (1999). Globalization and the synchronization of transport terminals. *Journal of Transport Geography* 7: 255-261.
Rodrigues, V.S., Stantchev, D., Potter, A. and Naim, M. (2008). Establishing a transport operation focused uncertainty model for the supply chain. *International Journal of Physical Distribution and Logistics Management* 38, 388-411.
Roso, V. (2007). Evaluation of the Dry Port concept from an environmental perspective: A note. *Transportation Research* Part D 12(7): 523-527.
Roso, V., Woxenius, J., Lumsden, K. (2009). The Dry Port Concept – Connecting Seaports with their Hinterland. *Journal of Transport Geography* 17: 338-345.
Roson, R., Soriani, S. (2000). Intermodality and the Changing Role of Nodes in Transport Networks. *Transportation Planning and Technology* 23: 199-213.
Stank, T.P., Goldsby, T.J.(2000). A Framework for transportation decision making in an integrated supply chain. *Supply Chain Management* 5: 71-77.
Stahlbock, R., Voss, S. Operations research at container terminals: a literature update, *OR Spectrum* 30: 1-52.
Van Geenhuizen, M. (2000). Interconnectivity of Transport Networks: a Conceptual and Empirical Exploration. Transportation Planning and Technology 23: 199-213.
Vervest, P., Li, Z. (2009). The Network Experience – New value from smart business networks. Berlin: Springer.
Visser, J., Konings, R., Pielage, B.-J., Wiegmans, B. (2007). A new hinterland transport concept for the port of Rotterdam: organisational and/or technological challenges? Proceedings of the Transportation Research Forum, North Dakota State University.
Woxenius, J., Roso, V. and Lumsden K. (2004). The Dry Port Concept – Connecting Seaports with their Hinterland by Rail. In Licheng, S., Notteboom, T. (Eds.), Proceedings of the First International Conference on Logistics Strategy for Ports (pp. 305-319), Dalian Maritime University Press, Dalian (China).
Zuidwijk, R., Veenstra, A. (2010). The Value of Information in Container Transport: Leveraging the Triple Bottom Line. Working paper, Erasmus Research Institute for Management, Rotterdam.

FOR THE OPTIMAL ALLOCATION OF THREE MODES FOR THE INTERCONTINENTAL TRANSPORTATION

Takashi Amemiya[1*] and Katsuhisa Okita[2]

[1]Department of Business Administration and Information,
Faculty of Business Administration,
Setsunan University,
17-8, Ikedanakamachi, Neyagawa,
Osaka Prefecture, 725-7508 JAPAN
e-mail:amemiya@kjo.setsunan.ac.jp

[2] United Parcel Service Company
3-3-6, Ishida, Minato-ku, Osaka, 545-0033 JAPAN
e-mail: kokita@ups.com

Abstract: The existence of a worldwide trend in which the world's manufacturing site is moving to East Asia is well known. The products produced in these districts are transported to the advanced nations to meet their consumption demand. Among such advanced nations, the U.S.A. has the largest demand. It should be noted that these Asian districts used for the production sites are rather restricted. Therefore, the amount of the products transported from these restricted districts of Asia to the U.S.A. is becoming tremendously large and are causing very serious traffic problems. The new products are required to be transported swiftly by air. Once the consumption and market demand are stable, the products should be sent rather more slowly but in larger amounts. However, the airport capacity of China is quite restricted and the transport volume is becoming large. As a result, transport cost is increasing and also the time required for transport is increasing. Now a third method so-called Sea and Air transport is appearing. Its cost and time in transit takes a mean position between Air and Ocean freight service. However, there has been no reasonable strategy to allocate these three methods of transportation. Here this problem is considered under the assumption that all products are stored in the inventory and are transported following the demand from the consuming districts. The most efficient transport method is determined.

1. INTRODUCTION

As is shown in the previous paper[3], it is a well-known tendency thatmany advanced nations have established many firms in the far east for thecheap labor cost and transported their products to many worldconsuming countries. These consuming countries are usually developedcountries and the production sites are rather restricted. Therefore, theproducts are concentrated to rather small number of hub ports for theirtransportation. As these products, there appear many seasonal productssuch as textile products or Xmas gifts. Every year such kinds of productappear seasonally and their movements resume similar styles. Althoughsome products are not seasonal, such as cellular phone or some electronics devises, manufacturers produce new models rather cyclically and as a result their movements assume the same styles. We may consider that the cycle is usually 6-months long from experiences, although the cycle does not have special meanings. Generally the transportations of these products have particular characteristics. The new products should be sent as rapidly as possible after their production, because, the early appearance gives the new products large competitive advantages. For that reason the air transportation is usually used at this stage. However, since there are more products to be sent by air, air transportation isbecoming more difficult in both economically and in terms of volume of transportation. The location of production is rather concentrated in small area, and as a result, the airport to be used is also restricted. After the first period of such quick transportation, a slower but large amount of transportation is needed. For this transportation, the one by sea is the usual selection. Although the transportation by sea is slow, the cost of the transportation is very low and the amount to be sent at one time is extremely large. The difference of these two methods is extensively large both from the point of cost and from the point of the capacity. It is natural that the third method of medium speed and medium amount become necessary. This method is the so called sea-and-air transportation. Generally, this method means a combinational transportation. At first the products are sent by sea to some neighboring airport where the services are obtainable more easily and then they are sent by air as further transportation to the final destination. There exist ship services especially useful for such purposes. The most popular routes are seen the ones through Japan. The details of this situation are shown in the paper [3]. The cost for this new method is almost half and the time required is almost double of direct air. Introducing this method, however, casts a difficult problem of how much this method costs and moreover how much should be sent by this method. On this problem, only few works have been appeared[1][2]. The calculation of the optimal allocation of these methods is first considered in this series of research. In the previous paper[3], only the basic characteristic of the problem is considered. In this paper optimal allocation is considered on the more realistic assumption on the speed of the consumption is considered. This report is consisted as follows. In 2, the importance of this method in relation to Japan is specially presented. Section 3 is devoted to the several assumptions description of the system. In 4 the costs and obtainable values by three modes are considered and from them the final profit is calculated. Algorithms for calculation of the profit is shown in 5. The efficiency of the introduction of*S&A*-mode is considered in 6.The concluding remark is given in 7**.**

2. SYSTEM DESCRIPTION

In this section several characteristics of the system is presented first and then by making some assumption the way to find the optimal transportation is studied.

2.1 Characteristics of the System

It should be noted that these seasonal products should be transported to the consuming district as fast as possible once they are produced. They are new products and the swift appearance gives the products a great competitive advantages. After the first period, the products are sent in large quantities to respond to the steady demands of the consuming districts and there fore the value of them decreases. This is the first characteristics. To start the analysis the time variabl***et*** should be defined. It represents time in the destination.

Definition 1(Definitionof Time) In this report time t start when the first product arrives at the destination.

(***Assumption I***) The value of the product monotonically decreasesas time elapses. Then, it approaches to a fixed value. How rapidly thevalue decays and how high the first value is varies much depending on the characteristics of the products.

(***Assumption II***) The transportation cost decreases monotonically as the required time increases. The cost by air is about hundred times higher than those by sea. And the cost by sea-and-air is about half to one-third of the cost by air.

On the top of the cost of transportation, the cost for the inventory should be considered.

(***Assumption III***) Inventory cost is a linear function of length of storage time and the amount of products.

2.2 Stand Point of the Analysis

Here the basic assumption for the analysis is presented. The total system of the inter-modal transportation involves many enterprises, consisting a very complicated system. Usually production, transportation and retailing are conducted by different companies and therefore their advantages and disadvantages are conflicting each other. These situations make the analysis considerably complex. In the analysis a different stand point may leads to a quite different results. In this report, our stand point is that of retailers one. And it is first assumed that all the products are already stored in the sender's inventory. However, the transportation should be conducted by the usual ways. The problem, therefore, is as follows.

If the amount of consumption and speed of it is assumed first, then whatmode of transportation and when should it be performed to satisfy their consumption?

2.3 Modeling

For the analysis of this system, several parameters and functions areneeded to be given. For the sake of simplicity, the transportation byair, by sea- and-air, by sea are called ***A***-mode, ***SA***-mode, and ***S***-mode, respectively. Here, the following parameters are introduced

2.3.1 Parameters

s: Amount of the product sent by one time of transportation by ***A***-mode.

d: Unit time span. Basically this is day.

μs:Amount of product sent by one time transportation by ***SA***-mode.~~¥¥~~

νs:Amount of product sent by one time transportation by ***S***-mode.~~¥¥~~

T: Total inventory time. That means this is the time fromthe beginning until there exists no product in the inventory.

(note) The above parameters indicate that μ times of s are sent at one time by ***SA***-mode and ν times amount of s are sent by ***S***-mode. These situations are caused by adopting containers for the transportation by sea and not for air. For ***SA***-mode, CFS (container freight station) cargo are used

2.3.2 Functions

On the basis of these assumptions several nonlinear functions are needed for the mathematical analysis of the problem. First of all, as is indicated in the previous subsection, estimated total consumingamount and the esti mated consuming speed are to be given first. Theestimated total consuming function *GC(t)* is a function of *t*, indicating the among of the total consumption of the products until *t* from the beginning. This function is defined as

$$GC(t) = 1 - \exp(-\omega t)(\omega t + 1) \tag{1}$$

Figure 4 shows the schematic view of this function. This figure indicates a model of how the consumption progresses. The consumption begins from *0%* at *t = 0* and monotoniaclly increases until *100%* after about 6 months or so. Although it is an id ealistic m odel, it is u sual that th e ret ailersestimate the pr ogress of t heir sal es and mak e the plan o f transportationof th e pro duct. On th is curve, more precise explanations are gi ven i n 3. 3.4. The curve$f_v(t)$ or transportation cost curve $f_{tc}(t)$are also needed to be assumed. They are defined as follows.
$f_v(t)$: The value of one unit product at *t = 0* that means the transportation is started.
$f_{tc}(t)$: The cost of transportation of one unit product to transport after *t*of its transportation starts.
$f_{ic}(t)$: The cost of inventory of one unit for *t*after its storage begins.
Owing to the assumption of monotone decreasing property, we may choose exponential functions of time such as

$$f_v(t) = a \exp(-\alpha t) + e, \tag{2}$$

as a candidate of the value curve, where a, α, e are constants,decided by the characteristics of the product. In the case of the newmodels of goods, it is usually extremely high. That means α is rather large. This is also the case of products in sev erer c ompetitive co nditions. T he s peedy int roduction t o the m arket is strongly r equired. The v alue d ecreases monotonically and it approaches to certain fixed value *e*.Also a candidate of the function for the transportation cost is given as

$$f_{tc}(t) = b \exp(-\beta t) + h, \tag{3}$$

where b, β, h are also constant, decided by the characteristicsof the product.It is also determined by the size and other characteristics of the products. Of course, this curve does not meanthere exists every service on this curve. It should be noted that this$f_{tc}(t)$indicates a cost of one time service to transport unitproduct at *t*.
It should also be n oted that there exist only threeservices on this curve, that is ***A***-mode, ***SA***-mode and ***S***-mode. Inthe usual case, t he cost by ai r is about 100 times larger than that ofby sea and the time required for this transportation is 3 days by ai rand 7 days se a and a ir an d 30 days b y sea. T hat means *h* is comparatively s mall an d almo st 1/100 of *b*, although it may not be zero. This situation im plies that the co st cu rveapproaches ve ry n ear th e static value *h* aft er almost 30 days from the system started. Here for the future use, the following time interval constants are introduced.

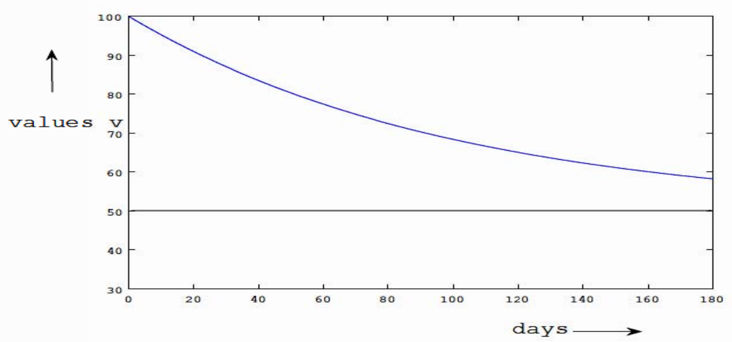


Figure 1: Schematic View of Value Function f_v

Definition 2 Time constantst_a, t_{sa} and t_s are defined as:
t_a is a required time for one time of ***A***-mode transportation.
t_{sa} is a required time for one time of ***SA***-mode transportation.
t_s is a required time for one time of ***S***-mode transportation.
Schematic views of these value function and cost function are given in Figure 1 and 2.
From the ob tained val ues, the i nventory cost shoul d also be subt racted. It is natural that th e i nventory cost is a lin ear function of time. Therefore it is given as

$$f_{ic}(t) = \gamma t. \tag{4}$$

Note that this is the one for a unit product.

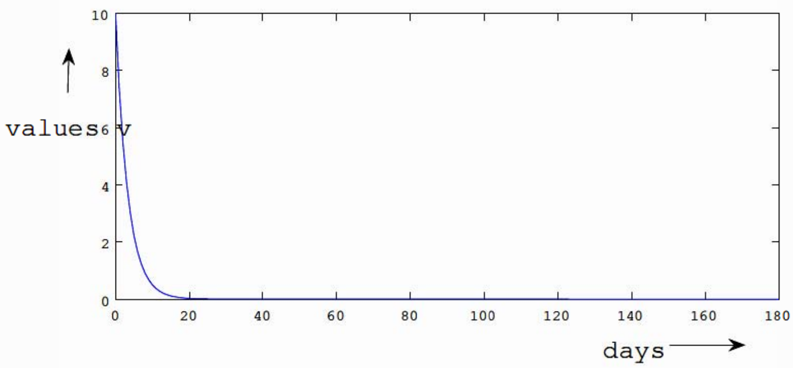


Figure 2: Schematic View of Cost Function f_{tc} (Transportation Cost)

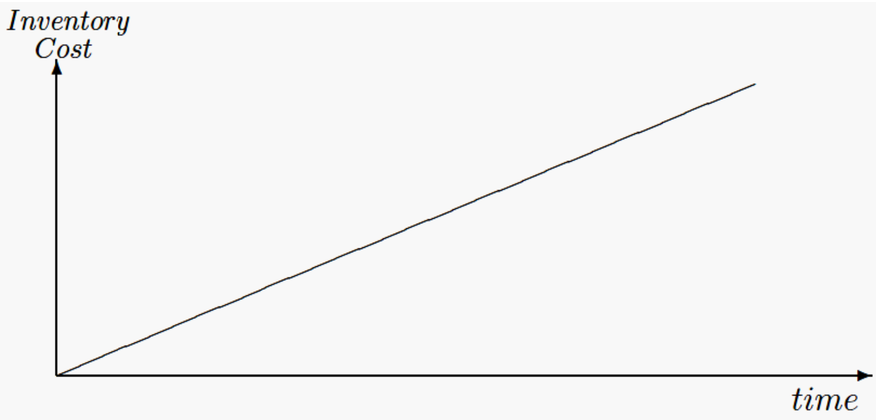


Figure 3: Schematic View of Inventory Cost Function f_{ic}

2.3.3 Assumptions on the Modeling

In the previous report the most basic situation for the transportation is considered. That means the products are basically transported directly after the production without saving in the inventory. More precisely, products are saved in the inventory only to wait until the proper amount for one of the three mode transportation is available. However, this does not necessarily reflect the actual circumstances. At present, production and transportation are conducted by different enterprises and therefore, all the products should be sent to the destination districts must be delivered at one time. There comes the new problem. In the consuming district there exists constant consuming demands and the total consuming amount is simply exhibited by the gross consuming function.

(***Assumption IV***) All the products to be sent are stored in the inventory at the beginning. The analysis starts when the first product arrives at the destination.

(***Assumption V***) The total consuming amount is a monotone increasing function of time. The total transported amount should not be allowed to come under this function at any time.

The most basic transportation system was analyzed in the previous report. In the report the transportation are performed soon after the production. In this report the production have already finished before the transportation. However, these both cases1 do not fully reflect the actual situation. The actual case are more complicated. These cases are to be analyzed in the future works.

2.3.3 Estimated Consuming Amount and Total Transportation Amount

Here the following two functions should be introduced. One is the function describing the total consuming function as the function of *t* and total transportation amount function. The total consuming function indicates the overall consuming amount from the beginning. The consumption begins from the start and increases and then comes to zero. The total consuming function is the integral of this amount of consumption and therefore it is the monotone increasing function and approaches to a certain fixed value. Here it is denoted as 100%.As a candidate of this function, the function given by (1) is considered. Figure 4 shows the schematic view of this function. This function is known as the function showing the response of a 2nd order control system to a unit step input function. It denotes how the total consumption reaches to 100% of the supplied products after the start of the consumption. ω is a parameter showing the speed of the response. In case ω is large the response becomes fast. Figure 6 shows the local view of this function around the start point.

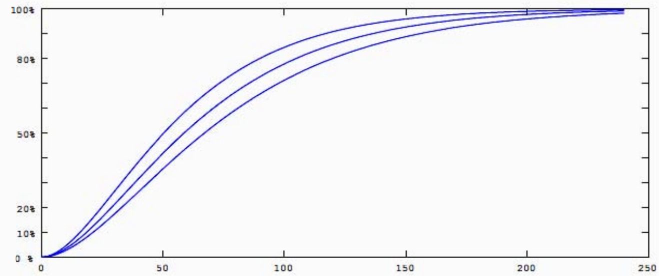


Figure 4: Schematic View of Total Consuming Function $GC(t)$

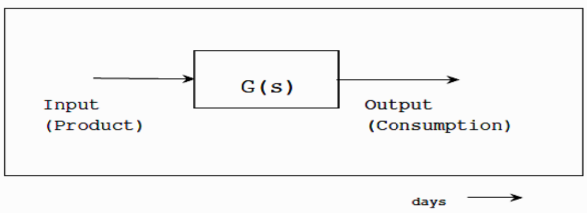


Figure 5: Transfer Function for Consumption

This curve is known as the indital response of a transfer function of so called 2nd order delayed system given as

$$G(s) = \frac{\omega^2}{(s+\omega)^2} \tag{5}$$

wheresis Laprace variable. It is conceived on the considerationthat the consumption may pick up speed after some initial hesitation. In this case the speed of consumption is given as

$$SGC(t) = \omega \exp(-\omega t)(\omega t + 1) - \omega \exp(-\omega t) \tag{6}$$

These curves are crude estimations of the actual consumptions. In theapplications of the theory, they should be constructed on the moreprecise and realistic analysis. The supply of the products should beaccomplished assuring these consumption. In Figure 7 the step functionsindicates the total transported amount. The first part was done by air and the second part is accomplished by sea-and-air and then the third part is done by sea.

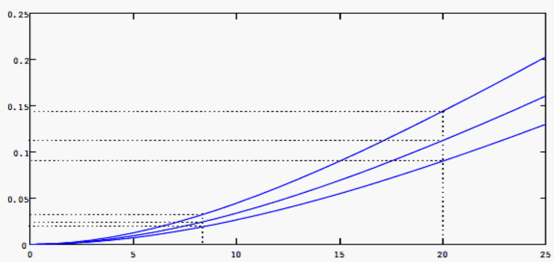


Figure 6: Schematic View of Total Consuming Function $GC(t)$ (Local View)

Theamount of the transportation by these three modes are quite different and the amount transported by air is a bout 1/100 of the amount by ***S***-mode.Since it is difficult to show the supply and consuming relation in one figure, the amount transported by ***A***-mode is neglected.

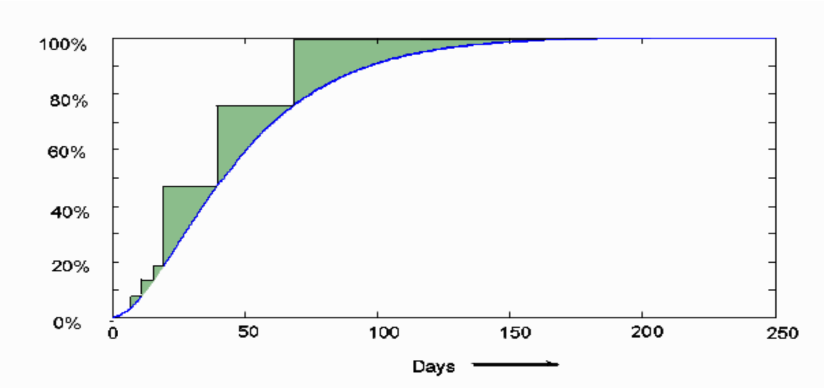


Figure 7 Schematic View of Transported Amount (Only ***SA***-mode and ***S***-mode are shown)

2.3.4 Calculation of the Obtained Value

The speed of the consuming function means the base of the obtainable values of the product. T he value of the product varies according to time. Therefore, the product of value and speed of the consumption implies the obtained value per day. Here the following function indicates this.

$$\begin{aligned} V_{obt}(t) &= SGC(t) * f_v \\ &= (\omega \exp(-\omega t)(\omega t + 1) - \omega \exp(-\omega t))(a \exp(-\alpha t) + e) \end{aligned} \quad (7)$$

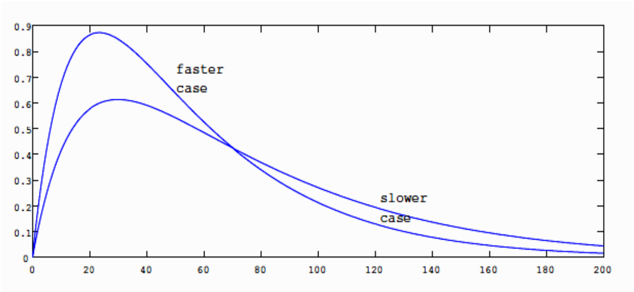


Figure 8 Schematic View of Obtainable Value

This figure shows that the obtainable values are very high for the early time of the transportation. That means the swift transportation is fairly efficient for the sales. Naturally, the swift transportation costs much. As was shown in the previous paper for one product

$$f_{tp}(t) = f_v(t) - f_{tc}(t) - f_{ic}(t) \quad (8)$$

shows the obtainable profit and the time point giving the peak of the curve is the best point for the transportation. O f course there exist only three services of the tran sportation. It can be safely said that th e nearest one becomes the most efficient service.
The final profit is obtained by subtracting the transportation cost from these obtainable values. Mathematically, the total obtained valuecan be calculated as the integral of the area between the obtainablevalue curve $V_{obt}(t)$ and the *x*-axis, that is

$$\text{Total Value} = TV = \int_0^\infty V_{obt}(t)dt \quad (9)$$

2.3.5 Calculation of the Cost

To calculate the genuin profit two kinds of costs should be subtracted. They are:
(1) Transportation Cost
(2) Inventory Cost
To mathematically define the problem, variables *x*, *y* and *z* are to be defined.

Definition 3
x denotes the number of transportation by ***A***-mode.
y denotes the number of transportation by ***SA***-mode.
z denotes the number of transportation by ***S***-mode.
The transportation cost by ***A***-mode, ***SA***-mode and ***S***-mode are denoted as $C_A(x)$, $C_{S\&A}(y)$ and $C_S(z)$, respectively. Th at means they are nonlinear functions of *x*, *y* and *z* respectively. The total cost *TC*is given as

$$TC = TC_A(x) + TC_{S\&A}(y) + TC_S(z) \quad (10)$$

As for the inventory cost, it is a linear function of both number of stored products and the length of the stored time.

$$IC = (\text{number of stored products}) * (\text{stored time}) \quad (11)$$

2.3.6 Constraint on the Transportation

The total consuming curve and consuming speed of the product are fixed,and some part of the transportation are to be automatically decided.Since the time required for the ***SA***-mode and ***S***-mode are at least 8 days and 20 days more than ***A***-mode, the products which are consumed during the first 8 days should be transported by ***A***-mode and those consumed during first 20 days should be transported by ***A***-mode or ***SA***-mode.

2.4 Problem Description

To start the problem the following restriction must be introduced.
(H1) Value curves and cost curve is fixed.
(H2) Inventory cost is a linear function of time.
(H3) The transportation should be restricted to three method.
(H4) Each transportation method has fixed transportation amountexcept the last one by sea.
(H5) The transportation amount of sea-and-air is μ-times larger than that of air.
(H6) Transportation amount of sea is again ν-times larger than that ofair.
The problem considered here is as follows:
Problem: Whether the ***SA***-mode is efficient to introduce. And if so, how much amount should be allocated to ***A***-mode ***SA***-mode and ***S***-mode and what are the point $t = \tau_{SA}$ and $t = \tau_S$ which make the profit maximum, where
$t = \tau_{SA}$is the time at which the storage by ***SA***-mode begins
$t = \tau_S$is the time at which the storage by ***S***-mode begins.
The above situation indicates that the transportation by air is performed from $t = 0$ to $t = \tau_{SA}$. Then from $t = \tau_{SA}$ to $t = \tau_S$the tran sportation by sea-and-air an d afte r $t = \tau_S$the tran sportation is only by sea. N ote that t hese ti me poi nts are thosein the destination, therefore to ensure these points, some modes of transportation must begins at the same time in the sending districts.
Then mathematically the problem can be shown as:
Mathematical description of the pr oblem Let the total profit obtained as a function of *x*,*y* and *z* by the transportation using three modes be denoted as $TP_{all}(x,y,z)$ Th en the problem should be as follows. Find *x*, *y*, *z* which satisfy

$$Max TP_{all}(x, y, z)$$

3. CALCULATION OF THE PROFIT

The total expected profit by the allocation of the three modes transportation is calculated here. It is assumed that x times***A***-mode tran sportation a nd *y* ti mes ***SA***-mode transportation and *z* ti mes ***S***-mode t ransportation a re p erformed. The i nventory cost and t ransportation cost are sa me for each t ime. Ho wever, the value function ch anges as ti me elapses. As is easily understood from the given figure, the product transported
by***SA***-mode arrives at th e d estination after τ_{SA} and the pr oducttransported by ***S***-mode arri ves at the de stination afte r τ_S.Therefore, b efore τ_{SA}only ***A***-mode s hould be adopted and before $\tau_S$only ***A***-mode o r ***SA***-mode can be used. T he

problem therefore is that after $\tau_A=0$ and before τ_{SA}, which is more efficient? ***A***-mode or ***SA***-mode? This is decided by performing the following calculation.

$$TP_{all}(x,y,z) = TV - TC_S - TC_{SA} - TC_S \tag{12}$$

3.1 Cost of the A-mode Transportation

By the assumption H5, the transportation amount is same at each time and is equal to s. Let time points t_{A1}, t_{A2},…,t_{Ax} be defined as

$$GC(t_{Ak}) = ks. \tag{13}$$

They are defined as the time at which the total consumption is equal to the total transported volume. Since the total consuming function is a nonlinear function, these points do not make fixed intervals. Terefore

$$C_A = TC_A + IC_A \tag{14}$$

and they are calculated as

$$\begin{aligned} TC_A &= xsf_{tc}(t_a) \\ IC_A &= \int_0^{t_{A1}} (a - GC(t))\gamma dt + \int_{t_{A1}}^{t_{A2}} (2a - GC(t))\gamma dt \cdots \\ &+ \int_{t_{A(x-1)}}^{t_{Ax}} (xa - GC(t))\gamma dt \end{aligned} \tag{15}$$

3.2 Cost by *SA*-mode Transportation

As in the previous sub-subsection, time points t_{SA1}, t_{SA2}, … t_{SAy} be defined as

$$GC(t_{SAk}) = xs + k\mu s. \tag{16}$$

Similarly to the previous subsection, the cost by ***SA***-mode included in the total transportation can be calculated as

$$C_{SA} = TC_{SA} + IC_{SA} \tag{17}$$

And they are calculated as

$$\begin{aligned} TC_{SA} &= \mu xsf_{tc}(t_{sa}) \\ IC_{SA} &= \int_{t_{Ax}}^{t_{SA1}} (xs + \mu s - GC(t))\gamma dt + \int_{t_{SA1}}^{t_{SA2}} (xs + 2\mu s - GC(t))\gamma dt \\ &+ \int_{t_{SA(y-1)}}^{t_{SAy}} (xs + y\mu s - GC(t))\gamma dt \end{aligned} \tag{18}$$

3.3 Cost by *S*-mode Transportation

After ***A***-mode and ***SA***-mode transportation, ***S***-mode transportation is performed. Following the previous sub-subsection, define time point t_{S1}, t_{S2}, …,t_{Sz} as

$$GC(t_{Sk}) = xs + y\mu s + k\nu s. \tag{19}$$

Total cost of ***S***-mode is given as

$$C_S = TC_S + IC_S. \tag{20}$$

And they are given as

$$\begin{aligned} TC_S &= \nu z s f_{tc}(t_s) \\ IC_S &= \int_{t_{SAy}}^{t_{S1}} (xs + \mu s - GC(t))\gamma dt + \int_{t_{S1}}^{t_{S2}} (xs + \mu s - GC(t))\gamma dt \cdots \\ &+ \int_{t_{S(z-1)}}^{T} (xs + \mu s - GC(t))\gamma dt \end{aligned} \quad (21\ )$$

4. TOTAL PROFIT OF THE PRODUCT

The total profit $TP_{all}$is obtained by substituting (9), (14), (17) and (20) into (12).

5. CONCLUDIG REMARKS

By the calculation of these values, it is shown that the introduction of SA-mode is fairly useful for the intercontinental product. The efficiency of the introduction varies depending on the products

Session G3: Seaport and Transportation 7

·Day2: Sep. 16 (Thu.)

·Time: 13:00 - 14:20

·Chair: Joachim R. Daduna

·Room: Iris, 4F

LOGMS

The 1st International Conference on Logistics and Maritime Systems

September 15th–17th, 2010 Busan, Korea

Lin

OPTIMIZATION OF INLAND CONTAINER TRANSPORTATION WITH AND WITHOUT CONTAINER SHARING

Herbert Kopfer[1], Jörn Schönberger[1], Sebastian Sterzik[1], and Ruiyou Zhang[2]

[1]Faculty of Business Studies & Economics, Chair of Logistics
University of Bremen
Wilhelm-HerbstStraße 5
Bremen, Germany
e-mail: kopfer@uni-bremen.de

[2]Institute of Systems Engineering
Northeastern University
#3-11Wenhua Road, Heping District
Shenyang, China
e-mail: zhangruiyou@ise.neu.edu.cn

Abstract: In this contribution we optimize the transportation of containers in the hinterland of a local area with one terminal and one depot for empty containers and trucks. There are several customers who want to receive goods by inbound containers and several customers who want to ship goods by outbound containers. Additionally, there are empty inbound containers and empty outbound containers. We present two different models corresponding to different scenarios for the transportation processes performed by a homogeneous and limited set of available trucks. In the first scenario (distinct container problem) empty containers are exclusively used by their owners and therefore must be sent to their predefined destinations. In the second scenario (shared container problem) empty containers can be interchanged among several owners and therefore can be arbitrarily used. By comparing the model for the distinct container problem with the model representing the shared container problem the benefit of container sharing can be analyzed.

1. INTRODUCTION

We present a truck and container scheduling problem and we model this problem for ahinterland transportation scenario with full and empty containers which are transported in a local area. A trucking company with a set of homogeneous trucks and a pool of empty containers is considered. We assume that there is a depot in which empty containers can be stacked and where the trucks are stationed. Additionally, it is assumed that there is a terminal in the areawhich can be a maritime port or a railway hub station to which trucks transport full and empty containers from customers' places and vice versa. In general, it would be possible to model the above situation for several depots and several terminals in the local area. But in order to keep the problem formulation simple we only present models for the *Inland Container Transportation Problem* (ICT problem) with one depot and one terminal. For further simplification werestrict our considerations to 40-foot containers. The ICT problem has been presented in Zhang *et al.*(2010). In contrast to Zhang *et al.*(2010) we consider in this paper two different versions of the ICT problem, one first version without allowing container sharing and a second version with the permission of container sharing. In the first version containers must be used for their predefined transportation task. In the second version containers can be arbitrarily interchanged in order to achieve improved solutions.

There are two types of containers, *inbound* and *outbound*containers. The containers located at the terminal that need to be moved to their destination (to the depots or their receivers) are called *inbound* containers. Reversely, the containers located at the depot or customers' places that need to be delivered to the terminal are called *outbound* containers. Moreover, each type of containers can be divided into full and empty containers. Thus, there are four types of containers demanding for transportation tasks that the company should carry out: inbound full, outbound full, inbound empty and outbound empty containers. First, an inbound full container has arrived from outside to the local area and is initially located at the terminal. It must be picked up by a truck at the terminal during a given terminal time window, must be delivered to its receiver (customer), and must be dropped off there. After being dropped off, the container is available at the customer location and is ready for being unpacked by the customer. When the inbound full container is completely handled at its destination, we obtain an empty container and a time window given for picking up the container at its current location. We have to move it to a depot or another alternative location by a truck.Secondly, an outbound full container is actually some freight that has to be transported in a container and is located at a customer's place. Thus, we should transport an empty container to the customer's location and deliver it during a given customer delivery time window which has been agreed on with the customer before. This empty container will be packed with freight by the customer. When the container is ready for shipment it can be picked up during a predefined customer pickup time window. It then has to bedelivered to the specified terminal during a predefined terminal time window. Of course, the before mentioned customer pickup time window must be consistent with the terminal time window for this container. Thirdly, an inbound empty container is also initially located at a terminal and is available to be picked up during its specific terminal time window. We should pick it up at the terminal and transport it to a depot or another alternative location regarding the time window of the chosen location. Finally, an outbound empty container means that we should pickup an empty container at the depot and deliver it to the specified terminal during the specified terminal time window. The topic of this paper is the optimization of the container flows in the local area for a given time peri-odas well as the resource planning and scheduling for a set of vehicles used for the needed container movements. For the analysis of the ICT we use an objective function which minimizes the total operating times of all trucks.

2. PROBLEM DESCRIPTION

For all full containers the origin (pickup location) and the destination (delivery location) is fixed by the problem data since these locations are defined by the required flows of goods carried in these containers. In the first scenario considered in this paper we assume that empty containers cannot be interchanged, maybe, since they have different owners and have to be used for their specific purpose or, maybe, since they have to reach their specific destination. This scenario is called *distinct container problem* during this paper. In the scenario of the distinct container problem the usage of empty containers being available at some location is determined in advance. That is why the origins and the destinations of all containers (empty containers as well as full containers) are fixed by the given data of a problem instance. In this case the optimization model related to the ICT problem comes up to a pickup-and-delivery problem with time windows (PDPTW) with each container movement representing a full truckload request for the PDPTW. The only difference to a usual PDPTW is that each customer has two time windows, one first time window for the delivery of a (full or empty) container in order to make the container available for the customer's loading or unloadingoperation and another second time window for picking up the container after the container has completely been handled by the customer.

But if it is allowed to interchange empty containers then we will have more flexibility. In this case, we can use any available empty container for any transportation. Throughout this paper this scenario will be called *shared container problem*. For the shared container problem, the decision which empty container will be assigned to the usage of which freight transportation task constitutes an optimization problem of its own. There are threetypes of empty containers which are available for the assignment to upcoming transportation tasks. The first type of available empty containers originates from the company's depot. The second type consists of all inbound empty containers located at the terminal. Finally, the thirdtype of available empty containers is constituted by all containers that have been emptied at a customer location and that are currently disposable for a new task.Available empty containerscan be used for two types of tasks. They can either be used as an empty outbound container (to be delivered to the terminal) or as a container which will be used to fulfill a customer's request for an empty container in the local area (i.e. the container will be packed with freight by this customer before it is transported to the terminal).Moreover, there is the opportunity for the trucking company to move the available empty containers to its depot. When empty containers can be interchanged, the origin of outbound empty containers and the destination of inbound empty containers are not defined by the problem data. The determination of these locations (i.e. a part of the input data of a PDPTW) is part of an optimization process itself. That is why the shared container problem cannot be modeled and solved as a usual PDPTW.

In this paper we discuss three approaches for modeling, describing and solving the ICT problem. The first approach refers to the distinct container problem, i.e. the model of the first approach describes the ICT problem without the possibility of container sharing. It turns out to be a PDPTW with a set of given container movements between customers, the terminal, and the depot. At the depot there are no time windows. For each container passing the terminal we have to respect its specific terminal time window. Each full container (inbound as well as outbound) has two time windows at its customer location (one for delivery and one for pickup).

The second and third approaches discussed in this paper refer to the shared container problem. The second approach is based on a sequential process for solving the two sub-problems of the ICT. The third approach pursues a simultaneous procedure for the solution of the ICT.

The second approach consists in the following two steps for solving the ICT. In the first step an optimal decision on the assignment of available empty containers to upcoming transportation tasks is aspired, i.e. in the first step it is tried to install minimum flows of empty containers in the local area in order to keep the total transportation demand of containers in the area as low as possible. The objective function used for the determination of the container flows is the minimization of the sum of the length of all distances that containers have to be transported. Of course, the determination of the container flows fixes an origin and a destination for each empty container which has to be transported. I.e., at the end of the first step, we have to solve the same type of problem as we have in the situation for the distinct container problem. That is why the model for the distinct container problem could also be used for the second step of the second approach. Since the flow of containers has to be performed by transportation processes fulfilled by the own fleet of the trucking company, the container flows have to be installed in that way that the given maximum number of coevally used trucks is not exceeded. The problem of minimum container flow with the important restriction of a limited resource capacity for transporting these containers is very interesting in general but it is not easy to solve. Since the problem in the first step of the second approach needs further investigation, this approach will not be pursued in the remainder of this paper.

Following the third approach we will solve the two sub-problems of the second approach in one single step, i.e. solving the assignment problem of empty containers simultaneously with the vehicle routing and scheduling problem induced by the originally given problem data and the compulsory assignment decisions.Using the model for solving the problem determines: a) where to deliver the empty containers released after inbound full/empty loads, b) where to pick up theempty containers for outbound full/empty loads, and c) in which order and by which truck the loads should be carried out.

3. EXAMPLE FOR CONTAINER INLAND TRANSPORTATION

Figures 1and 2 show a very small example for the ICT problem. Customers are illustrated by circles; the depot is illustrated by a rectangle and the terminal by a triangle. The flow of goods is shown in Figure 1. There are two customers represented by the nodes 1 and 2. The flow of goods from customer 1 to the terminal is demonstrated by the arc g_1 and the flow from the terminal to customer 2 is demonstrated by g_2. The flow of goods is made possible by means of containers. The time window for the availability of a container at the customer's location iis given by $[s_i,e_i]$. Additionally, there are terminal delivery time windows for outbound containers and terminal pickup time windows for inbound containers. Customer 1 will have to pack the container provided to him during the time window $[s_1,e_1]$. The container of customer 1 has to reach the terminal respecting the terminal delivery time window for this container and will then leave the local area via the terminal. Customer 2 will receive a container carrying a flow of goods shown by the arc g_2. For this customer the time window for unloading the container is $[s_2,e_2]$.

Figure 1. Flow of goods in the local area

Figure 2. Distinct container problem

The flow of goods induces a flow of containers. Figure 2(a) shows the flow of containers for the case that the containers used for customer 1 and 2 cannot be interchanged (i.e. the situation of the distinct container problem). The container C_1respectively C_2will be used for the realization of the flow of goods g_1 respectively g_2. The flow of the empty container C_1 is denoted by C_1(E) and afterwards when this container is loaded at customer site 1 its flow as a full container is denoted by C_1(F). The flow of the full container C_2 fromthe terminal to the customer 2 is shown as arc C_2(F) and after this container is unloaded by customer 2 its flow continues as an empty container to the depot on the arc denoted as C_2(E). As mentioned above there is an availability time window for containers at each customer's site. We assume that the customer delivery time window for a container to be delivered to customer i will be $[s_i–\varepsilon,s_i]$ and the customer pickup time window will be $[e_i,e_i+\varepsilon]$, respectively, with ε denoting the amount of time that a container may arrive earlier at a customer's site than necessary or the amount of time that the container is allowed to remain at a customer's site after the availability time window is over.

The flow of container requires corresponding truck operations. Figure 2(b) shows the transportation processes needed to implement the intended container flows. The solid lines illustrate the transport of containers by a truck and the dotted lines illustrate truck movements without any container. The bold solid lines indicate the transportation of a full container while the semi-bold lines indicate the transportation of empty containers. The solid lines are marked by a denotation, for instance $OF(C_1,CW_2,TW_1)$. This denotation is used for describing the type of container, the identity of the container, and the relevant time windows. The first two characters denote the type of the container transported on that line:OE for Outbound Empty, OF for Outbound Full, IE for Inbound Empty, IF for Inbound Full. The first parameter within brackets identifies the container to be transported, e.g. C_1 for Container 1. The second parameter identifies the time window to be met when picking up the container. The values of that parameter might be CW_1respectively CW_2 for the first respectively the second time window of the customer location where the container has to be picked up. Alternatively the value of the second parameter might be TW_j for the time window which is relevant for container j at the terminal. Finally, the value of the second parameter might be "-" indicating that no time window is relevant for the pickup operation. The third parameter identifies the time window to be met for the delivery of the container at its destination. The possible values of the third parameter are the same as the ones for the second parameter. The dotted lines used for the illustration of empty container movements are marked by a denotation which describes the time windows for the locations at the origin and destination of that movement, for instance $(-,CW_2)$ for a truck movement from the depot to a customer who has to be reached at his second time window. The first parameter identifies the time window at the starting point of that empty truck movement and the second parameter identifies the time window at the endpoint of that movement. The values for the time windows of empty movements can be the same as for the time windows for container movements on the solid lines. Figure 2(b)demonstrates the case that the time windows and the limitation of available trucks do not allow any bundling or concatenation of transport processes to common tours. For this case Figure 2(b)shows all transportation processes which are necessary in the local area to fulfill the container flows shown in Figure 2(a). There are 10 transportation processes neededfor the transportation of the two containers. For each move of a container to or from the depot there will be needed a pendulum tour (i.e. 4 truck movements for the two containers). And for each move of a container between a customer location and the terminal there will be a tour with three transportation legs (i.e. 6 truck movements for 2 containers).

Figure 3. Shared container problem

The optimization model for the distinct container problem will minimize the transportation effort (in driving distances or operating times of the available trucks) for a given set of container movements. The two approaches for the shared container problem try additionally to minimize the container flows. Provided that the availability time windows [s_1,e_1] and [s_2,e_2] of the customers 1 and 2 allow that the same container can be used for both customers, the container flow illustrated in Figure 2(a) can be reduced to the container flow shown in Figure 3(a). As shown in Figure 3(b) the set of needed transportation processes will also be reduced.

4. MODELING THE CONTAINER INLAND TRANSPORTATION PROBLEM

The two models for the first and second approach use the following variables, parameters and constants.

$s(0) = \{0\}$	Start node (Depot)
$e(0) = \{3d + e + f + 1\}$	End node (Depot)
$P^i = (1, \dots, p)$	Pickup nodes (outbound full customers, first time window)
$P^o = (d + 1, \dots, d + p)$	Pickup nodes (outbound full customers, second time window)
$P = P^i \cup P^o$	
$D^i = (p + 1, \dots, d)$	Delivery nodes (inbound full customers, first time window)
$D^o = (d + p + 1, \dots, 2d)$	Delivery nodes (inbound full, second time window)
$D = D^i \cup D^o$	
$H^{OF} = 2\mathrm{d} + 1, \dots, 2\mathrm{d} + \mathrm{p}$	Terminal nodes (belonging to the number of outbound full customers)
$H^{IF} = 2d + p + 1, \dots, 3d$	Terminal nodes (belonging to the number of inbound full customers)
$H^{IE} = 3d + 1, \dots, 3d + e$	Terminal nodes (belonging to the number of inbound empty containers)
$H^{IF} = 3d + e + 1, \dots, 3d + e + f$	Terminal nodes (belonging to the number of outbound empty containers)
$H = H^{OF} \cup H^{OE} \cup H^{IF} \cup H^{IE}$	
$V = (s(0) \cup P \cup D \cup H \cup e(0))$	All nodes
K=1,…,m	Vehicles
$C^{IF} = d - p$	Inbound full containers
$C^{IE} = d - p + 1, \dots, d - p + e$	Inbound empty containers
$C^a = d - p + e + 1, \dots, d - p + e + 1 + z$	Additional empty containers (originating from the depot)
$C = C^{IF} \cup C^{IE} \cup C^a$	

e	Number of inbound empty containers
f	Number of outbound empty containers
z	Number of additional empty containers
m	Number of trucks
W	Waiting timefor the loading/unloading operation at a pickup/delivery node
M	Sufficiently big constant
t_{ij}	Required time for a truck to drive from node i to j
s_i/e_i	Time window of node i (i.e. TW_i for terminal time window of container i and CW_1 respectively CW_2 for the first respectively second customer time window)

Decision variables:	x_{ijk}	1, if truck k drives from node$i$to j; 0 otherwise
	y_{ijc}	1, if container c is moved from node $i$to j; 0 otherwise
	T_{ik}	Represents the starting time of truck k from node i
	L_{ic}	Represents the starting time of container c from node i

For a comprehensive survey of the different types of node setsFigure 4 illustrates their interrelations within the distinct and the shared container problem. Customers$i \in P$providing outbound full containersare defined as pickup customers. Additionally, customers $j \in D$ who receive inbound full containers from the terminal are declared as delivery customers. To constitute the first and the second time window of the customer locations each customer is represented by two vertices $(P^i \wedge P^o; D^i \wedge D^o)$. The terminal has to handle all types of containers and thus is split into four node sets$(H^{OF} \cup H^{OE} \cup H^{IF} \cup H^{IE})$. The additional possibilities of the shared container problem to allocate the container between the node sets are illustrated through the arrows and lay the basis for the following mathematical models.

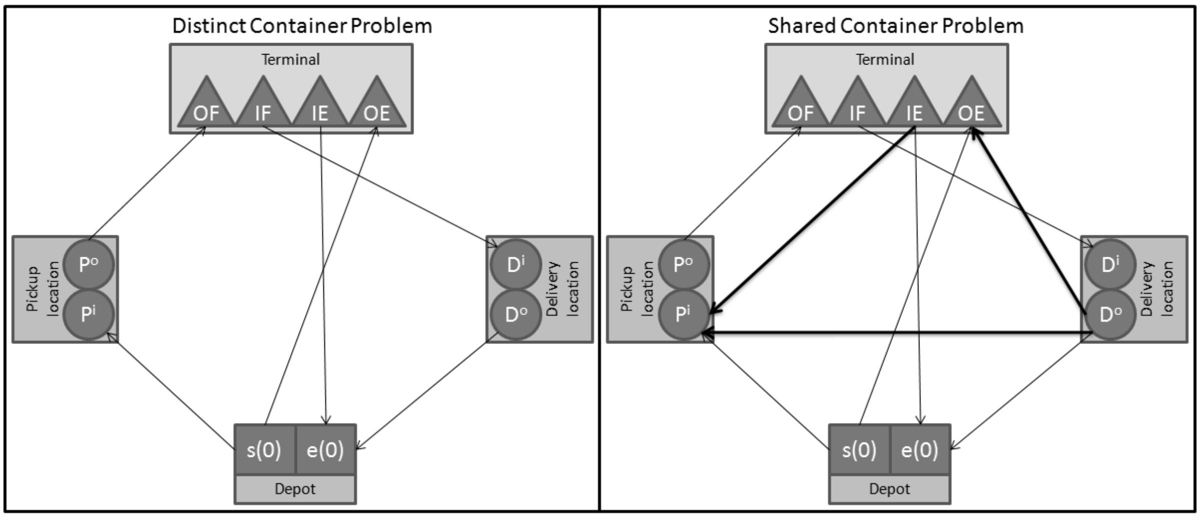


Figure 4. Possible container arcs between the node sets

The optimization model for the distinct container problem with the objective of minimizing the total operating times of all trucks consists in the equation (1) and the restrictions (2) to (14).

Objective function:	$minimze\ \ z = \sum_{k \in K}(T_{e(0)k} - T_{s(0)k})$		(1)
Restrictions:	$\sum_{j \in V}\sum_{k \in K} x_{ijk} = 1$	$i \in P \cup D \cup H; i \neq j$	(2)
	$\sum_{j \in V \backslash \{s(0)\}} x_{s(0)jk} = 1$	$k \in K$	(3)
	$\sum_{i \in V} x_{ie(0)k} = 1$	$k \in K$	(4)
	$\sum_{j \in V} x_{jik} - \sum_{j \in V} x_{ijk} = 0$	$i \in P \cup D \cup H; k \in K; i \neq j$	(5)
	$\sum_{k \in K} x_{i(d+i)k} = 1$	$i \in P^o$	(6)
	$\sum_{k \in K} x_{i(i-2d)k} = 1$	$i \in H^{IF}$	(7)
	$\sum_{k \in K} x_{s(0)jk} = 1$	$j \in P^i \cup H^{OE}$	(8)
	$\sum_{k \in K} x_{ie(0)k} = 1$	$i \in D^o \cup\ H^{IE}$	(9)
	$T_{jk} \geq T_{ik} + t_{ij} - M(1 - x_{ijk})$	$i,j \in V; k \in K$	(10)
	$T_{ik} + t_{i(d+i)} + W \leq T_{(d+i)k}$	$i \in P^i \cup D^i; k \in K$	(11)
	$T_{ik} + t_{i(d+i)} + W \leq T_{(d+i)m}$	$i \in P^i \cup D^i; k, m \in K$	(12)
	$s_i \leq T_{ik} \leq be_i$	$i \in V; k \in K$	(13)
	$x_{ijk} \in \{0,1\}$	$i,j \in V; k \in K$	(14)

The objective function (1) deals with the minimization of the total operating time of the trucks. Restriction (2) assures that every vertex is visited exactly once. While (5) guarantees the route continuity, restrictions (3) and (4) mean that a truck starts a tour from the depot and beyond ends the tour at this location. The following two restrictions ensure that a truck which picks up an outbound full container from a pickup vertex(i.e. during the second time window) drives to the terminal. Furthermore, a truck driving to aninbound full terminalnode has to move an inbound full container to the related delivery customer. Constraints (8) and (9) guarantee that every pickup locationand every outbound empty terminalnode is supplied by empty containers from the depot.Furthermore empty containers from the terminal or from a delivery customer node must be moved to the depot. Whilethe time continuity duringa tour is assured by (10),the waiting time for the loading and unloading operation at the pickup and delivery customer vertices is guaranteed through (11). According to this restriction, constraint (12) states that the container is transported as well from the pickup or delivery node after waiting time, if two different trucks pass the first and the second time window of the same customer location. Finally, (13) assures that a truck reaches a location in its defined time window.

The optimization problem for the shared container problem is represented by the objective function (15) and the restrictions (16) to (44).

Objective function: $minimze\quad z = \sum_{k\in K}(T_{e(0)k} - T_{s(0)k})$ (15)

Restrictions :

$\sum_{j\in V}\sum_{c\in C} y_{ijc} = 1$	$i \in P \cup D; i \neq j$	(16)
$\sum_{i\in V}\sum_{c\in C} y_{ijc} = 1$	$j \in H^{OF} \cup H^{OE}; i \neq j$	(17)
$\sum_{j\in V}\sum_{c\in C} y_{ijc} = 1$	$i \in H^{IF} \cup H^{IE}; i \neq j$	(18)
$\sum_{i\in H^{IF}}\sum_{j\in V} y_{i(i-2d)c} = 1$	$c \in C^{IF}$	(19)
$\sum_{j\in e(0)\cup D^i}\sum_{c\in C^{IE}} y_{ijc} = 1$	$i \in H^{IE}$	(20)
$\sum_{j\in V} y_{s(0)jc} = 1$	$c \in C^a$	(21)
$\sum_{j\in V} y_{s(0)jc} = 0$	$c \in C^{IF} \cup C^{IE}$	(22)
$\sum_{i\in V} y_{is(0)c} = 0$	$c \in C$	(23)
$\sum_{j\in V}\sum_{c\in C} y_{ijc} = 0$	$i \in H^{OF} \cup H^{OE} \cup e(0); i \neq j$	(24)
$\sum_{i\in V}\sum_{j\in H^{OF}\cup H^{OE}\cup e(0)} y_{ijc} = 1$	$c \in C; i \neq j$	(25)
$\sum_{i\in s(0)\cup D^o}\sum_{c\in C} y_{ijc} = 1$	$j \in H^{OE}; i \neq j$	(26)
$\sum_{c\in C} y_{i(d+i)c} = 1$	$i \in P^i \cup P^o \cup D^i$	(27)
$\sum_{j\in V} y_{jic} - \sum_{j\in V} y_{ijc} = 0$	$i \in P \cup D; c \in C; i \neq j$	(28)
$L_{jc} \geq L_{ic} + t_{ij} - M(1 - y_{ijc})$	$i,j \in V; c \in C; i \neq j$	(29)
$L_{ic} + t_{i(d+i)} \leq L_{(d+i)c}$	$i \in P^o; c \in C$	(30)
$L_{(2d+i)c} + t_{(2d+i)i} \leq L_{ic}$	$i \in D^i; c \in C$	(31)
$L_{ic} + t_{i(d+i)} + W \leq L_{(d+i)c}$	$i \in P^i \cup D^i; c \in C$	(32)
$s_i \leq L_{ic} \leq e_i$	$i \in V; c \in C$	(33)
$\sum_{j\in V}\sum_{k\in K} x_{ijk} = 1$	$i \in P \cup D \cup H; i \neq j$	(34)
$\sum_{j\in V} x_{s(0)jk} = 1$	$k \in K$	(35)
$\sum_{i\in V} x_{ie(0)k} = 1$	$k \in K$	(36)
$\sum_{j\in V} x_{jik} - \sum_{j\in V} x_{ijk} = 0$	$i \in P \cup D \cup H; k \in K; i \neq j$	(37)
$\sum_{k\in K} x_{ijk} \geq y_{ijc}$	$i \in s(0) \cup P^o \cup D^o \cup H^{IE} \cup H^{IF}$; $j \in V; c \in C; i \neq j$	(38)
$\sum_{k\in K} x_{ijk} \geq y_{ijc}$	$i \in V; j \in P^i \cup D^i \cup H^{OF} \cup H^{OE}$; $c \in C; i \neq j$	(39)
$T_{jk} \geq T_{ik} + t_{ij} - M(1 - x_{ijk})$	$i,j \in V; k \in K; i \neq j$	(40)
$T_{ik} = L_{ic}$	$i \in P \cup D \cup H; k \in K; c \in C$	(41)
$s_i \leq T_{ik} \leq e_i$	$i \in s(0) \cup e(0); k \in K$	(42)
$x_{ijk} \in \{0,1\}$	$i,j \in V; k \in K$	(43)
$y_{ijc} \in \{0,1\}$	$i,j \in V; c \in C$	(44)

The objective function deals with the minimization of the total operating and waiting time of the trucks. While constraints (16) to (33) assure the containers' routes, (34) to (44) guarantee the routes of the trucks. Thus, (16) states that every pickup and delivery node is visited once by a container. Restrictions (17) and (18) assure that the terminalis either visited or left once according to the export or import vertices. The starting and end nodes of the different kinds of containers are guaranteed by restrictions (19) to (26). Thereby IF containers need to be moved from the terminal to the delivery customers.While IE containers can be transported to ingoing pickups nodes or the depot, (21) states that additional empty containers originate from the depot.The three types of containers are not allowed to start their tour from a different starting node stated by (22)-(24). Constraints (25) and (26) assure that the containers will end their tour either at the depot or the export terminal. While route continuity is stated by (28) restriction, (27) ensures that a container visits the related node of a pickup, terminal or delivery node and furthermore passes the filling and emptying process of the container. According to these restrictions the time continuity ((29-31))and the service time for the loading and unloading operation at the pickup and delivery nodes ((32)) have to be held. Finally, restriction (33) assures that a container reaches a location in its defined time window.The truck constraints (34)-(44) are comparable to the restrictions for the distinct container problem. Attention should be paid to restrictions (38), (39) and (41) which assure that the trucks are interlinked with the containers and pass every location at the same time. Hence, the trucks cover the containers' routes but can skip the waiting time at the pickup and delivery vertices.

5. CONCLUDING REMARKS ON FUTURE RESEARCH

The models presented in Section 4 deliver a precise formulation of the ICT problem with and without container sharing. The problem is interesting from a theoretical point of view since there are two levels of transportation planning which are intervolved with each other and have to be matched together. On the lower level there is the resource planning for the containers which are to be used for transportation, and on the upper level there is the resource planning for the trucks that are needed for the movement of the containers which are planned for transportation. CPLEX has been used for the solution of small instances of models presented in Section 4. By comparing the results of the solutionsof instances of the distinct container problem with the results obtained by the solutionsof instances of the shared container problem the benefit of container sharing can be estimated and analyzed in dependence of the characteristics of the given problem instances. This benefit is measured with respect to the reduction of the transportation costs for inland container transportation in a local area. Unfortunately, CPLEX is only able to solve small problem instances of the ICT problem. We assume that the benefit reached by container sharing is relatively small for undersized problem instances and will grow tremendously when the problem instances become bigger and bigger. In order to check this assumption in our future work we will develop heuristic approaches for the solution of the ICT problem for both scenarios, with and without container sharing.

ACKNOWLEDGEMENTS

This research was supported by the German Research Foundation (DFG) as part of the Collaborative Research Centre 637 “Autonomous Cooperating Logistic Processes – A Paradigm Shift and its Limitations” (Subproject B9).

6. REFERENCES

Zhang, R.; Yun, W. Y. and Kopfer, H.(2010).Heuristic-Based Truck Scheduling for Inland Container Transportation.*OR Spectrum (ORS), Special Issue on "IT-based Planning and Control of Seaport Container Terminals and Freight TransportationSystems*", *32*:787-808.

SHORT SEA SHIPPING AND RIVER-SEA SHIPPING IN THE MULTI-MODAL TRANSPORT OF CONTAINERS

Joachim R. Daduna

University of Applied Business Administration at Berlin
Badensche Str. 50 - 51
D - 10825 Berlin, Germany
e-mail: daduna@hwr-berlin.de

Abstract: The constantly increasing quantitative and qualitative requirements for the terrestrial container and Ro /Ro transport can not only be dealt with in road and rail freight transport and from transportation on inland waterways in the upcoming years. Therefore, suitable solutions have to be found which include other modes of transport, whereas both economic and ecological factors as well as macroeconomic considerations are of importance. One possible approach is to increase the use of Short Sea Shipping and River-Sea Shipping so far less applied. The underlying structures here are presented and reviewed for their advantages and disadvantages. Potential demand structures are identified and illustrated by various examples. The paper concludes with analysis and evaluation of these concepts and the summary of necessary measures for their implementation.

1. POLITICAL FRAMEWORK FOR FREIGHT TRANSPORT

The realization of cargo traffic, both as inland and port hinterland transport, largely occurs in road freight transport at the moment. This situation is very contradictory to the (worldwide increasingly coming to the fore) transport policy objectives, which provide a sustainable change of modal splits for the benefit of rail freight transport and freight transport on inland waterways. Considerations regarding the efficient use of resources and the reduction of mobility based pollution receive priority here. However the results of a realization of these goals should not be overestimated, as at this point the question has to be asked, to which extend a modal shift can actually be achieved under the existing technical and organizational framework and the requirements for operational processes in logistics (see e.g. Daduna 2009). This does not mean that there generally should not be measures made to shift from road transport to other modes of transport, but rather that the existing potentials must be exhausted, especially in long-distance haulage.

Targeted governmental measures in various countries, e.g. in the Federal Republic of Germany with the introduction of the road toll (for heavy trucks over 12 tones admissible gross vehicle weight on highways), which has caused an (administratively enforced) increase in cost of road transport, do not show the aspired effect (see e.g. Bulheller 2006; Bühler / Jochem 2008). Only the economical behavior of suppliers of services in the road transport has led to noticeable ecological effects, e.g. by the increasing use of vehicles with lower category pollutant (see e.g. BAG 2009: 19p).

The (often existing and wanted) political prioritization of multi-modal freight transport for rail / road has not yet led to the expected results regarding a significant change in modal split, as from user perspective in many cases process efficiency and adequate quality of services is not provided. In addition, in rail transport there are the often existing capacity restrictions regarding the available network infrastructure, as well as (for example within the *European Union* (EU)) the sometimes significant incompatibilities in cross-border traffic. This especially concerns the monitoring and control technology and the energy supply as well as with the legal framework (see e.g. Pachl 2004, S. 16ff).

Another possibility is the inclusion of inland waterway and maritime navigation in the structures of multi-modal transport, although there are (process-related) limits given. The inland waterway navigation can only offer limited bigger shift potentials because of the geographical structures of the available network of waterways. Also in maritime navigation accordant restrictions regarding the access possibilities to the (often also close to the customer and smaller) ports and therewith to the hinterland occur, for example with a (further) increase in ship sizes.

An increasingly discussed and also worldwide in various areas implemented solution offers the *Short Sea Shipping* (SSS) (also in the context of feeder traffic), where upon a larger number of smaller ports (with local / regional importance) is involved in the configuration of transport processes. In the focus of attention are multi-modal transport chains, in which primarily the (coastal) shipping is efficiently linked with the (classical) terrestrial modes of transport. A specific extension of these considerations results from the integration of *River-Sea Shipping* (RSS), because not only coastal transport routes are used here, but with a suitably designed inland network of waterways also access to the hinterland is given (see e.g. Charles 2008). The aim is a sustainable use of these transport concepts in the context of multi-modal transport chains, especially regarding the above-mentioned shift from (long-distance) road traffic to offers involving the SSS and the RSS.

Firstly, the transport policy framework of SSS and RSS is outlined and the basic (traffic) structures of operational processes within this market segment are introduced. This is followed by a consideration of the advantages and disadvantages from the economical and ecological as well as the macroeconomic point of view. Potential demand structures are identified and illustrated by various examples from different regions. The conclusion is a critical analysis and evaluation of the SSS and the RSS in view of the competition with mono-and multi-modal terrestrial transport.

2. BASIC STRUCTURES OF SHORT SEA SHIPPING

The SSS and the RSS underlying structures, which are not only part of logistic concepts developed in recent years, but were through the centuries (including inland waterway navigation) the indispensable backbone of national and international trade (and continue to be). But with the occurring of rail freight transport and later of road transport a large shift to these two terrestrial modes of transport take place (see e.g. Aldcroft 1996). An important reason there for was the attained independence of existing natural and artificial (water)traffic networks by these traffic related innovations, which finally formed an essential precondition for industrialization in Europe.

For transport policy considerations the aim is an even greater use of SSS and RSS especially in the EU (see e.g. COM EC 2001: 14p and 47pp; COM EC 2004; Frerich / Müller 2006: 197pp; COM EC 2009; Medda / Trujillo 2010), whereas an effect of the modal splits in the (intra-European) freight haulage at expense of road transport is the superior goal. Even if, based on the EU-inland transport (including the national transport), only the (classical) terrestrial modes of transport are considered, they only cover 60% of the total volume, i.e. 40% were handled in the SSS and RSS in recent years (see e.g. EC COM 2001: 7; Aberle 2009: 49pp).

Applied to maritime transport in the EU-27 (in 2007), 61% of the total emergence are related to the SSS (and the RSS) (see e.g. Amerini 2009), while in the Federal Republic of Germany the percentage is at 55.5% (partly because of the comparatively short coastline) (see e.g. Reim 2009). In view of the (unaccompanied or accompanied) multi-modal freight transport in its various forms it is shown that this (still) has a smaller impact. The percentage of container transport is at only 11.3%, which is still below the Ro/Ro transport with 13.5% (see e.g. Amerini 2009) resulting primarily from ferry transports. With a share of 67.6% the main focus is on tank shipping (48.1%) and the transport of dry bulk cargo (19.5%).

Based on the growing importance of the multi-modal transport (and especially the container traffic) in the context of globalized supply chains is (not only in the EU) a stronger focus on this transport segment is needed. Prerequisites therefore are service offers within the SSS (or in the RSS) geared to the demand structures, which may persist (even under cost effectiveness aspects) next to the competition. In addition, also environmental factors are of great importance (in conjunction with transport policy discussions) that must be seen during the use of the SSS and the FSS (see e.g. COM EC 2004; Baird 2007). The structures of the SSS can be divided into three basic forms, in which both mono-modal as well as multi-modal transportation operations can exist (see Figure 2.1).

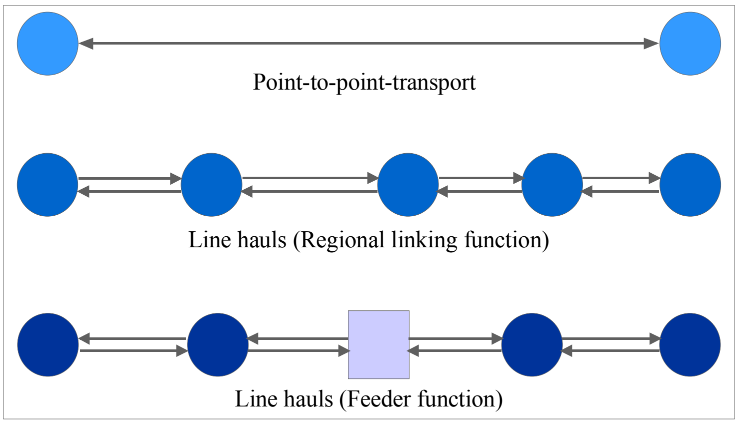


Figure 2.1. Basic structures of SSS

- *Point-to-point transport*: This may be for example the connection of two ports in multi-modal transport with adequate hinterland connections. Mono-modal connections between two economic zones and internal transport for companies with dislocated manufacturing sites are possible as well (with appropriate demand structures). In view of the temporal requirements (and hence the competitiveness as well) in transportation processes, Ro/Ro transport involving powerful (fast) vessel units (see e.g. Baird 2007) are in the foreground.

- *Line hauls (with regional linking function)*: These (often based on defined *Sea Motor Ways*) traffic can serve a regional linking of port locations, also in multi-modal transport including hinterland connections. Besides Ro/Ro traffic (see above) load unit transport (container, etc.) is of importance.
- *Line hauls* (*with feeder function*): In this case it is matter of feeder transport (with pick-up and delivery) for the regional linking of *Seaport Container Terminals* (SCT) with a (higher-ranking) hub function and locally oriented ports (with capacity and performance at a lower level). In this form of the SSS and the involved hinterland offers, container transport is important due to in this case underlying transcontinental (or intercontinental) transport chains. However, depending on the structuring of hinterland transport, a change-over could take place within the SCT to Ro/Ro transport.

With an integration of the RSS opportunities to improve the operation structures quantitatively and qualitatively arise. This also allows a direct integration of locations in the hinterland with point-to-point traffic, whereby the necessary transshipment operations for the connection from SSS to inland waterway transport are not necessary. In the context of line hauls suitable inland ports can be served, for example via dead-end connections, and therefore additional demand potentials can be developed. However, it has to be considered that cost savings are possible through the elimination of transshipment operations on the one hand, but on the other hand also (could) cause negative effects on costs, as in the RSS correspondingly smaller vehicle units with lower capacities and therefore higher costs (per unit) are used.

The structures outlined above show that for the configuration of offered services in the SSS and the RSS a standardization of services is possible only to a limited extend due to the complex requirements and restrictions. Within the service design a number of technical parameters have a significant role, which partly (can) set clear limits. This not only refers to the dimensions of vessels used, but for example also to port infrastructure (in terms of capacity and the operational efficiency) and the available hinterland connections as well as the existing inland waterway network (with inclusion of the RSS), especially in relation to the maximum rate of flow.

3. POSSIBILITIES AND LIMITATIONS OF SHORT SEA AND RIVER-SEA SHIPPING

The various (political) efforts in enhancing the use of the SSS and the RSS (see e.g. KOM EC 2004; Frerich / Müller 2006: 197ff; Baird 2007), not only at European level, clarify the expectations of this form of (international) freight transport. An important role here have, next to transport policy objectives, economic as well as environmental considerations. The focuses of attention are the following positive aspects (see e.g. Chang et al. 2007: 16p; Denisis 2009: 31ff; Medda / Trujillo 2010):

- *Improving energy efficiency*: Due to the high energy demand in the transport sector, but with strongly differing proportions of the various modes of transport (see e.g. Kamps 2003: 14p), the inclusion of the SSS and the RSS can be useful in multi-modal transport chains, since the lowest energy consumption (applying to traffic performance in tkm) occurs there.
- *Reduction in pollution*: Realized traffic shifts (so from road transport to the SSS and the RSS) lead to an increase in energy efficiency within the transport sector and thus inevitably also to a reduction of the road traffic generated pollutions (see e.g. Kamps 2003: 18). This effect is increased by tightened (technical) regulations in shipping (see e.g. IMO 1997; Frerich / Müller 2006: 302pp; Lee et al. 2010).
- *(Spatial) displacement of emissions*:
  In agglomerations with high emission load (pollutants, noise, etc.) partial relief can be achieved by modal shifts, but at which additional burdens arise in other places (but usually to a lesser extent).
- *Relief on road infrastructure*: Amongst other things positive effects arise from a reduction of the congestion generated by trucks, especially on highways, by what road safety can be increased.
- *Reduce costs for the maintenance of road infrastructure:* By reducing the number of (heavier) trucks, the wear of road infrastructure and therefore also repair costs can be reduced. Furthermore, network expansion and / or extension activities and thus investment costs can be avoided.
- *(Nationwide) extension of transport infrastructure*: Both with the SSS and with the RSS in some cases existing gaps in the transport infrastructure may be closed (see e.g. Sánchez / Wilmsmeier, 2009), i.e. sub-networks will be linked and necessary connections will be built (bridging effect).

Other positive effects may result within the multi-modal transport concerning costs, but only at larger distances (see e.g. Gather et al. 2008, p. 210ff; Aberle 2009: 550ff). The mode-specific fixed costs and the transshipment costs in the intersections at first for shorter distances (over) compensate the cost benefits of multi-modal transport chains. However, also various influences have to be seen that impact negatively on the performance of the SSS and the RSS. This primarily results from limitations relating to (logistical) fields of application (see e.g. Chang et al. 2007: 18p; Denisis 2009: 33ff):

- *Longer transport times*: Due to the (relatively) low rate of feed in maritime transport and the interruptions occurring within interfaces in multi-modal transport chains, transportation times can arise which do not meet customer requirements.
- *Aggregation / disaggregation of freight flows*: Time delays have to be considered in the interfaces between the different modes of transport, for example in terminals during transshipment from road transport to the SSS (and vice versa). These arise on the one hand (due to capacity aspects) from a collection of load units (container, swap bodies, etc.) and the other hand from the not at the same time available number of needed trucks for transportation.
- *Additional transshipment costs*: Changing modes of transport are not only a time problem, but they also generate additional (terminal) costs, i .e. it is necessary here to develop appropriate technical solutions eliminating this weakness of intermodal transport chains.
- *Administrative efforts*: With the increasing complexity of the (transport logistical) processes the coordination effort to monitor and to control the included modes of transport accordingly, especially regarding the time requirements of shippers.

Considering the advantages and disadvantages it is shown that the advantageousness of SSS and RSS which largely results from the economic and ecological effects in the logistic area is in general not given or only partially. I.e. it must be asked whether a sufficient acceptance can be achieved on the shippers' side. Key features include (see e.g. Paixão Casaca / Marlow 2005) the following requests on the user side (see also Medda / Trujillo 2010):

- Qualification and development of multi-modal transport network structures with respect to a capacity expansion, and a reduction in transport time.
- Sufficient frequency of service offerings and increase of reliability within transport operations.
- Improvement of efficiency and effectiveness of port facilities.
- Construction of company-wide information structures for better planning and monitoring of transport processes.
- Development of appropriate (cooperative) organizational forms involving all stakeholders.

For a target-oriented implementation a close cooperation is required between the (for the transport infrastructure) relevant governmental institutions and the operators of involved logistic facilities due to the different responsibilities. Only then the structural requirements for the organization of efficient operational processes in SSS or in RSS can be set up and thus a sustainable marketability can be achieved. Possible applications for the SSS and the RSS are outlined in the following section.

4. APPLICATIONS IN EUROPE

As a number of studies show (see e.g. Sánchez / Wilmsmeier 2005; Baird 2007; Brooks / Trifts 2008; Charles 2008; Perakis / Denisis 2008; Ng 2009) both the SSS and the RSS is used for freight transport worldwide in various forms, especially as an alternative to the (classical) terrestrial modes of transport and also as a completion to existing transport infrastructure. In addition, in different regions due to geographical structures considerable potentials for a (spatial) extension of these offerings arise, as well as for greater development of individual market segments, for example in container (and, in particular, empty container (s. u.a. Le-Griffin / Griffin 2010)) and Ro/Ro transport, unless there are no restrictions due to *market regulations* (s. u.a. Brooks / Frost 2004; Perakis / Denisis 2008; Medda / Trujillo 2010). In the following section, therefore, the situation in the EU (and the neighboring European area) is outlined exemplary, in view of the current use and the existing potentials.

To analyze and evaluate the SSS (as well as the RSS) within the EU and neighboring countries consistent (Europe-wide) structures cannot be assumed due to spatially differentiated conditions and requirements. Likewise, a consideration oriented towards the Member States does not make sense, since the potential demand (particularly in the SSS) here is one in larger scale cross-border freight transport with differing shares in type and scope (see e.g. Amerini 2009). For these reasons, a regionalization with the aim of adapting to the spatial requirements will be taken as a basis (see e.g. Amerini 2009). A division into five regions is assumed here (*Baltic*, *North Sea*, *Atlantic*, *Mediterranean* and *Black Sea*) which may not be regarded as enclosed spaces as they are (in a different extent) linked to each other.

The EU has a crucial role in view of the (largely) transnational transport procedures, within the European area, which is also increasingly recognized (see e.g. KOM EG 2009). A concretion of the (recent) implementation of the SSS within the EU is coming from the revised *Trans European Transport Network* (TEN-T)-structures (see e.g. Baird 2007). Destined is a network of ports via defined Sea Motor Ways (as central axes) (see Figure 4.1). Currently, four *Sea Motor Ways* are defined, the *Baltic*, *Atlantic*, *Western Mediterranean* and *Eastern Mediterranean* (see e.g. Psaraftis 2005; Baird 2007) These are integrated in the EU-wide traffic concept, i. e. they constitute an essential foundation for the organization of multi-modal transport chains in the container transport as well as particularly in the Ro/Ro transport (see e.g. Baird 2007; Styhre 2009) including the terrestrial traffic infrastructure (see e.g. Paixão Casaca / Marlow, 2009).

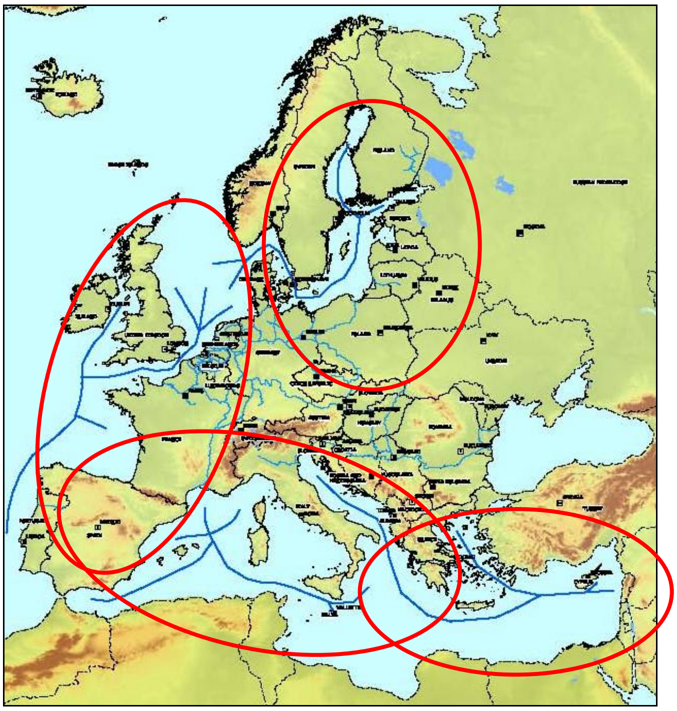


Figure 4.1. Sea Motor Ways in Europe

Independent of the issue of a consideration orientated on regional structures or transport axes, significant potentials for a modal shift and therewith for the SSS and the FSS are revealed in view of the economic and settlement structures within Europe and also within individual countries, (see e.g. Baird 2007). Transport technical basis here can be for example the not yet (in the SSS) enough used port locations, provided that they have sufficient hinterland connections (see e.g. Ng 2009) or have sufficient local demand potential. It is of importance for the Mediterranean (see e.g. Foschi et al. 2005), since the SSS is the only efficient link in freight transport between the islands and to the mainland.

The RSS also has gained in importance within the EU in various sections of the inland waterway network in recent years (see e.g. Charles 2008). Larger rivers, such as the Rhine, have an important role in this context (Fig. 4.2). However, the market share is still below 10% of the total transport volume of inland waterway navigation, caused, for example, by the low capacities of the relative large but mostly small sized network. Therefore the operating vessels are of smaller size with a reduced capacity. Significantly greater potentials arise in Eastern Europe, e.g. in the Russian Federation and in the Ukraine, where it is emanated from a share of over 30%.

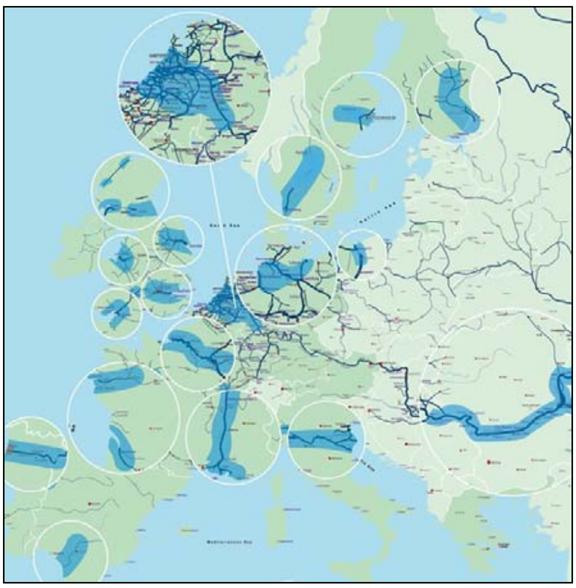


Figure 4.2. Potentials of RSS in inland waterways of the EU

5. ANALYSIS AND EVALUATION

The importance of SSS and RSS will be increasing (worldwide) in the long term, particularly in view of environmental aspects. Prerequisite for a sustainable implementation in the field of container and in Ro/Ro transport is not only a question of (transport) policy and technical framework and the existing potential of demand, but fundamentally a sufficient acceptance by users, i.e. essentially the shippers. Initial point for a critical evaluation in this case is a SWOT analysis (see e.g. Denisis 2009: 35f), whose outcome is outlined in Figure 5.1. It is distinguished here in *intra-company factors* (strengths / weaknesses) and *company external factors* (prospects and risks).

Strengths W	**eaknesses**
° Relatively high energy efficiency ° Low pollution ° Capacitive relief on road networks ° Reduction in expenditures for construction and maintenance of traffic infrastructure	° Low speed ° Integration of additional ports ° Additional handling costs (for container transport) ° High coordination costs (planning, control and monitoring of the load unit movements)
Prospects	**Risks**
° (Transport) policy prioritization of SSS and RSS ° Long-term growth in multi-modal freight transport ° Relief of central SCT ° Unfavorable cost developments in road transport (fuel, wages, etc.) ° Foundation of diagonal cooperations in transport and logistics industry	° Possible bottlenecks in the ports at rising of the SSS and the RSS ° Development of administrative cost structures (port taxes, bureaucracy, etc.) ° Prioritization and subsidization of competing modes of transport (in particular the rail freight transport)

Figure 5.1. SWOT analysis on SSS and RSS from a user perspective

Necessary condition for long-term acceptance is, primarily, a sufficient efficiency of services offered. However, it is taken into account (see e.g. Brooks / Trifts 2008) that identifiable cost benefits only generate a relocation effect if a sufficient quality (particularly regarding the timeliness and reliability) is perceived in the operational processes, i.e. there exists, as far as this is not given, a (partially) inelastic reaction. Only if an attractive offer within the multi-modal transport is placed on the market, with SSS and RSS as a basis, it will be possible to counteract effectively, particularly in competition with the (mono-modal) road freight transport.

The long-term objective, therefore, must be seen in a consistent development of differentiated (and client adapted) structures in the SSS (as well as in the RSS). This requires appropriate strategies (see e.g. Baird, 2007; Paixão Casada / Marlow 2009) and their implementation in a close networking of all companies involved in the (multi-modal) transport chains. In addition, the necessary framework has to be not only created, but also be developed continuously by the governmental authorities involved. Of crucial importance here is the planning (as well as the enforcement) of intermodal network structures and their financing, also in the context of (cooperative) *Public Private Partnership* (PPP) models (see e.g. Sánchez / Wilmsmeier 2005).

Only if it works to involve the SSS and the RSS in the multi-modal transport structures in a way, that in the long-term economically and ecologically sustainable processes in transport operations can be offered, they will enforce to compete for the customers in the market. The focus here is, as shown above, the relief of existing terrestrial transport infrastructure (see e.g. Baird , 2007) and an extension of inadequate network structures (in road and rail traffic), particularly in developing and emerging countries (see e.g. Sánchez / Wilmsmeier 2005). The geographic structures are given in many regions, so that a global transfer of experience can help to make more and more use of this (also from an ecological point of view) reasonable transportation offered.

5. REFERENCES

Aberle, G (2009). Transportwirtschaft. 5., überarb. u. erg. Aufl. Oldenbourg, München / Wien

Aldcroft, D.H. (1996). The eclipse of British coastal shipping, 1913 - 21. in: Armstrong, J. (ed.): *Coastal and short sea shipping.* Scolar Press, Aldershot: 163 - 177

Amerini, G. (2009). In 2007, EU-27 Short Sea Shipping continued growing but at a slower rate. in: *eurostat - Statistics in focus - Transport* 58/2009

Baird, A.J. (2007). The economics of Motorways of the Sea. *Maritime Policies & Management* 34: 287 - 310

Brooks, M.R. and Frost, J.D. (2004). Short sea shipping - A Canadian prespective. *Maritime Policies & Management* 31: 303 - 407

Brooks, M.R. and Trifts, V. (2008). Short sea shipping in North America - Understanding the requirements of Atlantic Canadian shippers. *Maritime Policies & Management* 35: 145 - 158

Bühler, G. and Jochem, P. (2008). Wen iger Em issionen v on Speditionen d urch Ver kehrsverlagerung - Wel che Potenziale haben politische Instrumente? *Internationales Verkehrswesen* 60: 278 - 281

Bulheller, M. (2006). Verlagerung auf die Schiene muss teuer erkauft werden. *Internationales Verkehrswesen* 58: 353 - 355

Bundesamt für Güterverkehr (BAG) (ed.) (2009). *Marktbeobachtung Güterverkehr - Jahresbericht* 2008. Köln

Chang, Y.-T., Lee, P.T .-W., Kim, H.-J., Shin, S.-H. and Kim, M.J. (2007). *Short sea shipping study - A report on successful SSS models that can improve ports´ efficiency and security while reducing congestion, fuel costs, and pollution*. Asia-Pacific Economic Cooperation (APEC) Transportation Working Group / Inha University

Charles, L. (2008). Sea-river shipping competitiveness and its geographical market area for the Rhône-Saône corridor. *Journal of Transport Geography* 16: 100 - 116

Denisis, A. (2009). *An economic feasibility study of short sea shipping including the estimation of externalities with fuzzy logic*. Diss., The University of Michigan

Foschi, A.D., Peraldi, X. and Rombaldi, M. (2005). *Inter-island links in Mediterranean short sea shipping networks*. Discussion Paper n. 52; Dipartimento di Scienze Economiche / Università di Pisa

Frerich, J. a nd M üller, G. (2006). *Europäische Verkehrspolitik - Band* 3: *Seeverkehrs- und Seehafenpolitik*, *Luftverkehrs- und Flughafenpolitik*, *Telekommunikations-*, *Medien und Postpolitik*. Oldenbourg, München / Wien

Gather, M., Kagermeier, A. and Lanzendorf, M. (2008). *Geographische Mobilitäts- und Verkehrsforschung*. Bornträger, Berlin / Stuttgart

International Maritime Organization (IMO) (1997). *International convention for the prevention of pollution from ships - Annex* VI: *Prevention of air pollution from ships*. (http:// ww.imo.org/conventions/mainframe.asp?topic_id=255)

Kommission der Europäischen Gemeinschaften (KOM EG) (2001). *Weissbuch - Die europäische Verkehrspolitik bis* 2010 - *Weichenstellungen für die Zukunft*. (KOM(2001) 370 endgültig)

Kommission der Europäischen Gemeinschaften (KOM EG) (2004). *Vorschlag für eine Verordnung des Europäischen Parlaments und des Rates über die Durchführung des zweiten* "*Marco Polo*" *Programms über die Gewährung von Finanzhilfen der Gemeinschaft zur Verbesserung der Umweltfreundlichkeit des Güterverkehrs* ("*Marco Polo* II"). (KOM(2004) 478 endgültig)

Kommission der Europäischen Gemeinschaften (KOM EG) (2009): *Mitteilung und Aktionsplan zur Errichtung eines europäischen Seeverkehrsraums ohne Grenzen*. (KOM(2009) 10 endgültig)

Le-Griffin, H.D. and Griffin, M. (2010). Managing empty container flows through short sea shipping and regional port systems. *International Journal of Shipping and Transport Logistics* 2(1): 59 – 75

Lee, P.T.-W., Hu, K.-C. and Chen, T. (2010): External cost of domestic container transportation - Short-sea shipping versus trucking in Taiwan. *Transport Reviews* 30: 315 - 335

Medda, F. and Truj illo, L. (201 0). Short sea sh ipping - An an alysis of its d eterminants. *Maritime Policies & Management* 37: 285 - 303

Ng, A.K.Y. (2009). Competitiveness of short sea shipping and the role of port - The case o f North Europe. *Maritime Policies & Management* 36: 337 - 352

Pachl, J. (2004). *Systemtechnik des Schienenverkehrs*. 4., überarb. u.erw. Aufl. Teubner, Stuttgart / Leipzig

Paixão Casaca, A.C. and Marlow, P.B. (2005). The competitiveness of short sea shipping in multimodal logistics supply chains - Service attributes. *Maritime Policy & Management* 32: 363 - 382

Paixão Casaca, A.C. and M arlow, P.B. (2 009). Log istics st rategies for sh ort sea sh ipping operating as part of multimodal transport chains. in: *Maritime Policy & Management* 36: 1 - 19

Perakis, A.N. and Denisis, A. (2008). A survey of short sea shipping and its prospects in the USA. *Maritime Policy & Management* 35, 591 - 614

Psaraftis, H.N. (2005). EU ports policy - Where do we go from here? *Maritime Economics & Logistics* 7, 73 - 82

Reim, U. (2007). Kombinierter Verkehr 2005 - Wachstum der Containertransporte in allen Verkehrsträgern. *Wirtschaft und Statistik* 59: 169 - 179

Sánchez, R.J. and W ilmsmeier, G. (2005). Short-sea shipping potentials in Central America to bridge infrastructural gaps. *Maritime Policy & Management* 32: 227 - 244

Sánchez, R.J. and W ilmsmeier, G. (2009). Mar itime secto r and po rts in th e ar ribbean - The case of C ARICOM countries. *United Nations CEPAL - Serie Recursos naturales e infraestructura* 140

Styhre, L. (2009). Strategies for capacity utilization in short sea shipping. *Maritime Economics & Logistics* 11: 418 - 437

Woitschützke, C.-P. (2006). *Verkehrsgeographie*. 3. Aufl. (Bildungsverlag EINS) Troisdorf

A CONFLICT-FREE SHORTEST-TIME ROUTING ALGORITHM FOR INTERMEDIATE-SCALE PATH NETWORK

Byung-Hyun Ha*, Tao Qin

Department of Industrial Engineering
Pusan National University
San 30, Jangjeon-dong, Geumjeong-gu
Busan 609-735, Republic of Korea
e-mail: bhha@pusan.ac.kr
qintao361@126.com

1. INTRODUCTION

A container terminal is a complicated system for transporting containers. It plays an important role in worldwide trade and it has become an essential intermodal interface between the sea transportation and the land transportation. Generally, a container terminal consists of quay area, transportation area, storage yard area, gate area and so on. There are vessels and quay cranes in the quay area. Yard cranes move on blocks of the yard area. Yard trucks travel among quay area, yard area and transportation area for loading or discharging containers. Containers in the yard area can be re-marshaled as well. External trucks delivery or receive containers among yard area, gate area and transportation area. An efficient containers transportation system is important for enhancing the performance of a container terminal due to the congestions and the travelling time of vehicles. However, it is difficult to resolve bottlenecks in the containers transportation system. There are many related optimizing problems, such as the design of a transportation layout, the shortest-time path finding for yard trucks etc. Hence, it is vital to develop some methodologies to resolve these problems in order to improve the productivity of the container terminal. Nevertheless, one of the most difficult problems is to plan and to control the paths of vehicles for transporting containers among different areas of the container terminal. A realistic transportation network of the container terminal consists of many different types of lane which allow the specific movements (e.g. forward movement, backward movement) of a vehicle. Many vehicles with different size and speeds move in the transportation network simultaneously. In addition, the collisions among all moving vehicles should be avoided and the efficiency of containers transportation should be kept as well. All of the issues above increase the difficulty of the vehicle routing problem in container terminal.

In our research, we focus on the conflict-free shortest-time path finding problem in an intermediate-scale network where the distance between nodes is shorter than the length of a vehicle but greater than the width of the vehicle typically. Moreover, the properties of the transportation network and the properties of vehicles in the container terminal are taken into consideration for developing an efficient algorithm. For example, the size of a vehicle, the direction of a vehicle and the moving directions of a vehicle are considered. We propose a formal model for developing an efficient algorithm. The definition of position is devised and the movement of position is given firstly. Free time windows and the time window graph are introduced for defining the conflict-free position and the conflict-free movement. In addition, the concepts of the space-reachability, the collision-free and the time-reachability are described. With the help of these definitions, the main procedure for finding a conflict-free shortest-time path is given based on the Dijkstra's algorithm. Moreover, the updating mechanisms for the time window graph are proposed as well. The time complexity of the proposed algorithm is analyzed and experiments are carried out to validate the performance of the proposed algorithm.

2. PROBLEM DESCRIPTION

The grid-based transportation networks can be classified into three different scale networks based on the distance between nodes. They are analyzed as follows.

Macro-scale network. In a macro-scale network, the distance between nodes is quite large. The vehicle can be treated as a point without mass and can occupy one node at a time. It is not necessary to consider the size of a vehicle, the direction of a vehicle, even the moving direction of a vehicle when generating a path.

Micro-scale network. A micro-scale network is the opposite of a macro-scale network. In the micro-scale network, the distance between nodes is quite short and a vehicle can occupy many nodes at a time. The size of a vehicle, the direction of a vehicle, and the moving direction of a vehicle are important when a vehicle is in the network.

Intermediate-scale network. An intermediate-scale network is the network of which the scale is between the macro-scale network and the micro-scale network. In the intermediate-scale network, the distance between nodes is shorter than the length of a vehicle but greater than the width of the vehicle. The vehicle can occupy several nodes at a time. Similarly, the properties of the vehicle are important for generating a path in the network.

When modeling for a same real transportation network with different scales, the number of nodes in the macro-scale network is less than in the micro-scale network typically. The relationship between a vehicle and its occupied nodes in the macro-scale network is simpler than in the micro-scale network. Hence, the efficiency of a path finding algorithm in the macro-scale network is higher than in the micro-scale network. Contrarily, because there are no enough nodes to describe the exact state of a vehicle in the macro-scale network, some properties of the vehicle are disregarded and the reality about the movement of the vehicle in the macro-scale network is lower than in the micro-scale network. The efficiency of the path finding algorithm and the reality about the movement of the vehicle in the intermediate-scale network is between the macro-scale network and the micro-scale network. Because we focus on the path finding algorithm for container terminal, both the reality about the movement of the vehicle and the efficiency of the path finding algorithm are important. However, it is hard to find an algorithm with both high reality and high efficiency. Hence a trade-off is made and the intermediate-scale network is chosen for the research of our path finding algorithm. In order to enhance the efficiency of the path finding algorithm, it is acceptable to sacrifice a part of reality of the vehicle movement without losing the generality.

2.1 Position, Network, and Vehicle

In the macro-scale network, a vehicle can occupy one node at most. Hence it is meaningless to define the representation of a position by using multiple nodes. In contrary, there are enough nodes to describe the position of a vehicle in the micro-scale network directly. Nevertheless, in the intermediate scale network, a reasonable representation of a position should be designed by considering the reality of the movement of the vehicle and the efficiency of the path finding algorithm.

Firstly, two definitions about the representation of a position are given. The real position represents the actual position of a vehicle while it moves in a real network. The logical position represents the position of a vehicle which will be applied to the path finding algorithm instead of the real position. Figure 2.3 shows two different models of the position for finding the conflict-free shortest-time path in the intermediate-scale network.

The properties of the network are listed as follows:

The network is not a bidirectional graph completely. Either in the vertical yard layout or the horizontal yard layout, the transportation network consists of different types of lane, such as unidirectional lanes, bidirectional lanes.

The distance between nodes is not the same. The distance between nodes could be either greater or smaller than the length of a vehicle. That is, the transportation network is irregular.

There are restrictions on the forward and backward movements of a vehicle. In the network, some lanes allow a vehicle moving both forwardly and backwardly, but other lanes in special regions only allow vehicle moving either forwardly or backwardly. The network becomes complicated and the difficulty of the path finding problem rises accordingly.

There is a major reason that this type of network is proposed for the conflict-free path finding of vehicle. If this type of network is proposed in the algorithm, the movement control of a vehicle on the operational level can be avoided. Because the movements of the vehicle are already restricted in the network based on the type of the lanes, it is not necessary to control the movement of vehicles in the operational level. Hence all kinds of costs can be reduced.

In our research, a vehicle is defined by the length, the speed and the center of the vehicle. The center of the vehicle is not the geographic center point of the vehicle but is defined as a point on the vehicle with an arbitrary position. A vehicle could occupy more than one node at a time while moving or waiting in the defined network. Hence, a set of nodes are proposed to establish the logical position of the vehicle. l_v^F and l_v^B are denoted as the lengths from the center of a vehicle v to its head and tail, respectively.

2.4 Assumptions

There are mainly four assumptions on the conflict-free shortest-time path finding problem. In our research, the first is about the speed of a vehicle. The second is about changing the moving direction of a vehicle. The third is about the logical position rule of a vehicle. The last is about the state of a vehicle at the source position or the destination position.

Each vehicle has the unique forward moving speed and backward moving speed. The former speed is greater than the latter one typically. Assume that the speed of the vehicle is constant while it is moving from one node to another. The vehicle can only stop for waiting while the center of the vehicle is on a node. That is, the speed of the vehicle can be 0. It shows the speed of a vehicle can be one of three values based on the moving direction and the position of the center of the vehicle in the network. s_v^F and s_v^B are the forward moving speed and the backward moving speed of the vehicle *v*. The vehicle only can reverse its moving direction while its center is on a node. It is prohibited to reverse the moving direction of the vehicle when the center of the vehicle is not on a node. As mentioned before, the logical position of a vehicle is different from its real position. The logical position will be applied to our conflict-free shortest-time path finding algorithm. Assume that a vehicle could be bent once at most and the vehicle can move forwardly or backwardly when its body is either straight or bent. When a vehicle stays at a source position or destination position, its body must be straight because of loading or discharging containers at that time.

2.4 Problem definition

Similarly to the problem defined by Kim and Tanchoco (1991), the conflict-free shortest-time path finding problem in this thesis is to find out the path starting from a source position and arriving at a destination position as soon as possible, without any collision to other vehicles and disrupting other active travel schedules, by considering the properties of the network and vehicles in an intermediate scale network. Figure 1 shows an example of the problem. Figure 2.12 shows that several vehicles move in an intermediate-scale network which consists of a set of nodes and a set of lanes. Vehicle A and vehicle B are already scheduled and a part of their paths are shown as dash lines. Our target is to find out a path for the vehicle C starting from its source position and arriving at the destination position as soon as possible, without any collision to other vehicles and disrupting any other active travel schedule, by considering the size, the moving speeds of the vehicle C. In summary, based on the model of the position and the assumptions, an efficient conflict-free shortest-time path finding algorithm will be proposed in an intermediate-scale transportation network by considering the properties of the network and vehicles for container terminal.

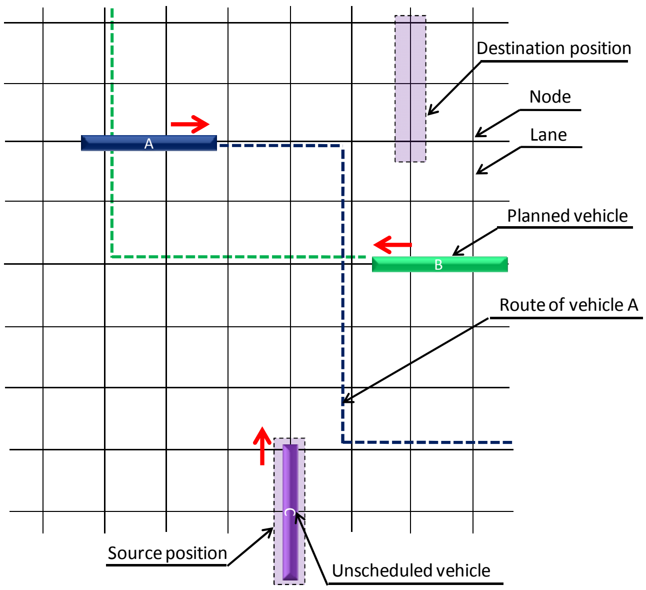


Figure 1. An example of scheduling vehicles in a network

3. ROUTING MODEL

In this section, a formal model is proposed to help define the path finding algorithm. First of all, the formulas about the distance in a network are given. Then the position and the movement of a vehicle are devised precisely. Next, the time window graph is introduced. The conflict-free position and the conflict-free movement are defined in detail. Finally, the formulas about the space-reachability, the collision-free, and the time-reachability are given.

We defined a transportation network *G* which consists of a set of nodes *N*, a set of forward moving arcs A^F, and a set of backward moving arcs A^B, denoted as: $G = (N, A^F, A^B)$, where $A^F \subseteq N^2$ and $A^B \subseteq N^2$. In a network which consists of a set of nodes and a set of lanes, two nodes are connected if there is one lane at least between them. Given two nodes n_1 and n_2 in the network *G*, n_1 and n_2 are *connected* if and only if $(n_1, n_2) \in A^F \cup A^B$ or $(n_2, n_1) \in A^F \cup A^B$. If n_1 and n_2 are *connected*, there should be a distance between them. The distance between nodes n_1 and n_2 can be denoted by $d(n_1,n_2)$ and $d(n_1,n_2) = d(n_2,n_1)$. In addition, if there is only one node n_1, the distance of this node is regarded as 0, denoted by $d(n_1)$.

In an intermediate-scale network, the distance between two nodes could be shorter than the length of a vehicle. That is, a vehicle can occupy more than one node at a time. Hence, a set of nodes is proposed to define the position of a vehicle. There are two sequences of nodes, $X = (x_1, x_2,\dots, x_n)$ and $Y = (y_1, y_2,\dots, y_m)$, where x_i and y_j are nodes of the network. $x_1, x_2, \dots, x_{n-1}$ are the nodes which are occupied by a vehicle from its center to the head where x_1 is the node which the center of the vehicle is located on. x_n is an unoccupied node which is nearest to the head of the vehicle, called as the guiding node. It can be used for guiding the forward moving direction of the vehicle. Similarly, $y_1, y_2, \dots, y_{m-1}$ are the nodes which are occupied by the vehicle from its center to the tail of the vehicle. y_1 is the same node with x_1 which the center of vehicle is located on. y_m is the guiding node which is nearest to the tail of the vehicle and can be used for guiding the backward moving direction of the vehicle. Furthermore, the center of a position is defined as the node which the center of the vehicle is located on. The guiding node is needed for modeling a position because it can not only help define the position precisely but also can declare the forward and backward moving directions of a vehicle. There is an example to explain the importance of the guiding node of a position.

In a network, a vehicle can move from a position to another forwardly or backwardly. Hence there are two types of movement, the forward movement and the backward movement. Given a vehicle v, for the sequences of nodes $X = (x_1, x_2,\dots, x_n)$, $Y = (y_1, y_2,\dots, y_m)$, $X' = (x_1', x_2',\dots, x_{n'}')$ and $Y' = (y_1', y_2',\dots, y_{m'}')$, v is *movable* from $P = (X, Y)$ to $P' = (X', Y')$, denoted by $P \rightarrow P'$. The center of the position P is the node which the center of the vehicle is on. For simplifying the explanation of the formal model, some notations are derived as follows. $\triangle(P, n)$ denotes the deviation between the center of the position P and node n. $\triangle(P, P')$ denotes the deviation between the center of the position P and the center of the position P'.

The method of the time window graph is efficient and effective to solve the conflict-free shortest-time path finding problem for container terminal. Two types of time window are introduced: the reserved time window and the free time window. The reserved time window is the time interval or the time window between a pair of entry and exit times of a vehicle at a node which is exclusively reserved by the vehicle. The free time window is the time interval between the reserved time windows which are not scheduled for vehicles, denoted by f'. The node related to the free time window f is denoted by f^N. The starting time and the ending time of the free time window f are denoted by f^A and f^B, respectively. The conflict-free position is defined as follows: given a vehicle v, for the sequences of free time windows $Z = (z_1, z_2, \dots, z_k)$, $W = (w_1, w_2, \dots, w_l)$ and nodes x, y, $\mathscr{P} = (Z, x, W, y)$ is a *conflict-free position*.

As mentioned before, a conflict-free position consists of a set of free time windows and two guiding nodes. A position can be made up of all nodes related to the free time windows and the guiding nodes. The conflict-free position based on a specific position highly depends on the time moment on the position. Figure 3.8 (b) shows that $\mathscr{P}_1$ with the time 10 is a conflict-free position but $\mathscr{P}_2$ with the time 20 is not a conflict-free position even thought the positions of two conflict-free positions are same. Given a conflict-free position $\mathscr{P}$, the earliest arrival time of $\mathscr{P}$ is defined as the earliest time that the center of a vehicle arrives at the node which is the center of the position $\mathscr{P}^P$, denoted by α. Given two conflict-free positions $\mathscr{P}$ and $\mathscr{P}'$, $\mathscr{P}'$ with α' (the earliest arrival time of $\mathscr{P}'$) is *conflict-free-movable* from $\mathscr{P}$ with α, denoted by $(\mathscr{P},\alpha) \Rightarrow (\mathscr{P}',\alpha')$.

ACKNOWLEDGEMENTS

This work was supported by the Grant of the Korean Ministry of Education, Science and Technology (The Regional Core Research Program/Institute of Logistics Information Technology).

5. REFERENCES

Kim, C.W. and Tanchoco, J.M.A. (1991). Conflict-free shortest-time bi-directional AGV routing. *International Journal of Production Research, 29:* 2377-2391.

OPERATIVE TRANSPORTATION PLANNING IN CONSUMER GOODS SUPPLY CHAINS

Hans-Otto Guenther[1*]

[1] Technische Universität Berlin
Germany
e-mail: Hans-Otto.Guenther@tu-berlin.de

Abstract: Transportation management in today's consumer goods industry can be characterized by a high proportion of outsourced transportation services. Due to rising freight costs consumer goods manufacturers are looking for opportunities to increase the efficiency of their transportation network. This study presents an operational transportation planning problem typical of the consumer goods industry focusing on a network of suppliers, production facilities and warehousing locations. It comprises an analysis of freight costs in a consumer goods transportation network based on the freight rate structures. In this analysis a number of opportunities for efficiency gains are identified and consolidated in an operative transportation planning problem which is then numerically investigated. Furthermore, an overview of processes and organizational structures in transportation management is given with special focus on the integration of existing commercial

Session H1: SCM 7

·**Day2: Sep. 16 (Thu.)**

·**Time: 14:40 - 16:00**

·**Chair: Ho Gyun Kim**

·**Room: Camellia, 5F**

SUPPLY CHAIN CHAMPION IN THAILAND: A CASE OF RUBBER TO TIRE INDUSTRY

Duangpun Kritchanchai[1*] Tuanjai Somboonwiwat[2] and Wirachchaya Chanpuypetch[3]

[1]Department of Industrial Engineering
Mahidol University
25/25 Phuttamonthon 4 Rd, Phuttamonthon, Salaya
Nakhonpathom, Thailand
e-mail: egdkc@mahidol.ac.th

[2]Department of Industrial Engineering
King Mongkut's University of Technology Thonburi
91 Pracha-utid Rd, Bangmod, Thungkru
Bangkok, Thailand
e-mail: tuanjai.som@kmutt.ac.th

[3]Centre of Logistics Management
Mahidol University
25/25 Phuttamonthon 4 Rd, Phuttamonthon, Salaya
Nakhonpathom, Thailand
e-mail: wirachchaya.c@egmu-research.org

Abstract: Supply chain champion is one of the research directions in Thailand logistics and supply chain management. This study begins with selecting target industry by using Thailand Competitiveness Matrix (TCM). Within the potential industry groups from TCM, all industries are considered by exporting value, government support, and value creation. It is found that rubber industry is one of the highest exporting values with high demand in the world market. Thus, Thailand rubber industry is selected for further determining the product champion. We applied analytic hierarchy process (AHP) to weight the evaluation criteria and to rank several end-rubber product alternatives. The passenger car tire industry is found in the first rank. Then, flows of rubber products in the chain of passenger car tire in Thailand are analyzed. In this paper, rubber products flows from upstream to midstream are exhibited. The purpose of this paper is to identify capacity balance in the chain. Number of firms and location for rubber production at midstream plants are recommended. Forecasting for future productivity is also taken into account. More efficient flows and less logistics cost are presented with the new location of midstream manufacturing.

1. INTRODUCTION

Due to the trend of nationalization and globalization in the recent decades, the importance of logistics management has been growing in various areas (Tseng *et al*., 2005). The definition of logistics is a part of the supply chain process that plans, implements, and controls the efficient, effective forward and reverse flow and storage of good, services, and related information between the point of origin and the point of consumption in order to meet customers' requirements (Council of Logistics Management, 1991).

The direction of Thailand logistics research is to enhance the country's competitiveness. Three performance indicators are considered namely; cost reduction, efficiency improvement, and value creation. The logistics system improvement should help costs reduction and enable Thailand to compete effectively. In Thailand, this approach is now considered in the overall of micro and macro level. It has been recommended that the potential industry of Thailand must be determined in order to identify the champion. Afterward, logistics and supply chain management must be applied on the supply chain of the product champion.

This paper is under the project of "The Supply Chain Champion in Thailand". The main purpose is to identify the product champion for Thailand's competitiveness. The study starts with selecting the potential industry in Thailand. It is found that rubber industry is one of the potential industry groups in Thailand Competitiveness Matrix (TCM). Then, criteria are determined and the product champion of the group is identified by using Analytic Hierarchy Process (AHP) method. We then analyze the flows from upstream to midstream of the product champion. The directions of industry's competitiveness are provided in this study.

The paper is structured as follows. The analysis of Thailand's industries competitiveness is depicted in section 2. In this section, the criteria are determined for the product champion selection. In section 3, the current situations of rubber products are analyzed to define the problem in the chain. The direction for supply chain improvement will be proposed in section 4. Conclusions and further research are shown in section 5.

2. THAILAND'S INDUSTRY COMPETITIVENESS ANALYSIS

2.1 Industry groups in Thailand

Thailand Competitiveness Matrix (TCM) was created by National Economic and Social Development Board of Thailand. It is a model for comparing the international competitiveness of various industrial sectors and the attractiveness of different business types in the global market. Thailand Competitiveness Matrix is illustrated in Figure 1.

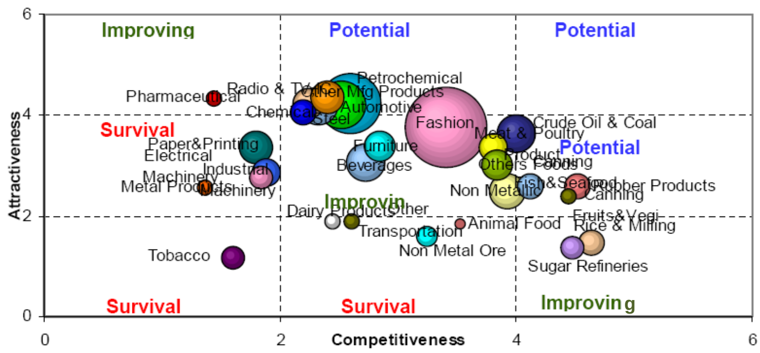


Figure 1. Thailand Competitiveness Matrix 2005: industrial sector
Source: Office of National Economic and Social Development Board

According to the TCM, Thai industries can be classified into three categories: potential industries, improving industries, and survival industries. All industries are considered by exporting value, government support, and value creation. Within the potential industry groups, rubber products, fish and seafood, and canning are more attractive and competitive than other groups. Being the number one rubber exporter in the world market, here we select rubber industry for the case study.

2.2 Thailand Rubber Industry and its Product Champion

2.2.1 <u>Thailand rubber industry</u>

Rubber industry in Thailand was established in 1900. Rubber production remained one of the country's major industries contributing to the growth of Thai economy. It is found that rubber industry is one of the highest exporting values with high demand in the global market. The major rubber exporting products of Thailand are tire and gloves. Exporting value are 31,642.62 million THB and 9,262.80 million THB in year 2009 respectively (The Office of Industrial Economics, Thailand, 2010).

Then we list down the top five products from these 2 groups. This will be considered as the alternatives for product champion evaluation. This is illustrated in Table 1.

Table 1. The top five export value of rubber products

Unit: Million THB

Rubber products	Year 2005	Year2006	Year 2007	Year 2008	Year 2009
Passenger car tire	29,093.06	31,462.72	33,623.50	35,419.25	31,642.62
Gloves	6,952.45	7,405.72	7,870.02	8,763.51	9,262.80
Motorcycle tire	3,543.66	4,268.57	4,409.24	4,847.05	4,353.62
Motorcycle inner tube	1,146.51	1,203.43	1,358.59	1,333.02	1,552.71
Bus/truck inner tube	873.93	869.60	876.27	937.75	970.38
Bicycle inner tube	448.55 484.	81 424.	85	-	-
Other tires	793.10	931.16	- -	-	
Tractor tire	1,328.55	1,340.91	- -	-	

Source: The Office of Industrial Economics, Thailand (2010)

From Table 1, the top five rubber products are selected for the alternatives for further identifying the product champion. The alternatives are passenger car tire, gloves, motorcycle tire, motorcycle inner tube, and bus/truck inner tube.

2.2.2 Product Champion for Rubber Industry in Thailand

The Analytic Hierarchy Process (AHP) is an approach that is suitable for dealing with complex systems. It is to provide choices from among several alternatives which provide comparison for the considered options. It is firstly proposed by Thomas L. Saaty in the 1970s (Bhushan and Rai, 2004). The AHP is based on the subdivision of the problem in a hierarchy form. The AHP helps organizing the rational analysis of the problem by dividing it into part. The analysis then provides the decision makers to make several pair-wise comparisons. The AHP can be used for analyzing different problems. It uses both qualitative and quantitative variables. The fundamental principle of the analysis is the possibility of connecting information based on knowledge to make decisions. In applying AHP, three steps are processed, the description of a complex decision problem as a hierarchy, the prioritization procedure and the calculation of result (Chauhan *et al*., 2008).

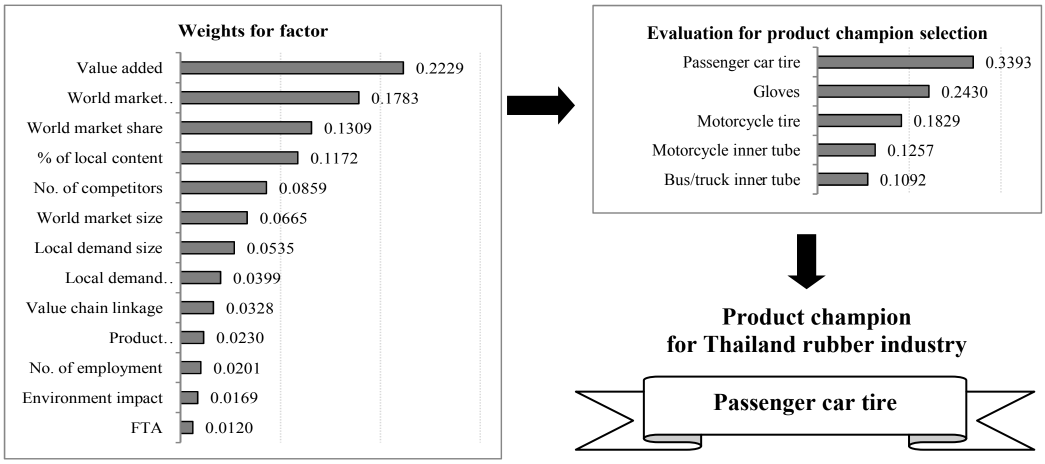
The first step in developing the AHP is to identify and then define the criteria. In this study, the criteria which has been used in Thailand Competitiveness Matrix are adopted. Group of criteria can be devided to five catagories namely factor inputs and production, global demand, domestic importance, degree of competition, and external effect. Then, thirteen criteria are selected from TCM and used for product champion selection. The criteria for Thailand product champion selection are shown in Table 2.

Table 2. The criteria for Thailand product champion selection

Group of criteria	Criteria	Definition of criteria
Factor inputs and production	% of local content	Percentage of local content usage for production
	Value added	Opportunity for value creation
	Value chain linkage	Opportunity for value chain linkage
	No. of Employment	Number of employment in the industry
Global demand	World market share	Share of the product in world market
	World market size	Market size of the product in world market
	World market growth	The growth rate in world market
Domestic importance	Local demand growth	The growth rate in local market
	Local demand size	Demand in local market
Degree of competition	number of competitors	Number of competitors in the market
	Product Differentiation	the uniqueness of products
	Free Trade Agreement (FTA)	Benefit from Free Trade Agreement (FTA)
External effect	Environmental impact	Impact of industry on environment

The alternatives and criteria are defined. We then develop a decision hierarchy. The first level of the model expresses the overall goal of this study which is to select product champion of rubber industry in Thailand. This goal can be achieved by considering the criteria in the second level. The second level presents thirteen criteria of the industrial performance. The third level shows five alternative products of rubber industry in Thailand.

The relative importance of criteria are determined through the pair-wise comparison method in AHP. We use the nine-point (1-9) numerical scale of AHP. The importance of the criteria were evaluated by group of expert. The relative weights of the criteria were next used to assess the alternatives of rubber product. The analysis aims to identify the ranking of the criteria. Based on the responses on the evaluation, the finding shows that the top three criteria are value added, world market, and world market share. Value added is given the highest weight of 0.2229 while world market and world market share is 0.1783 and 0.1309, respectively. Then, the results in rank of rubber products order are: passenger car tire (0.3393), gloves (0.2430), motorcycle tire (0.1829), motorcycle inner tube (0.1257), and bus/truck inner tube (0.1092). Criteria rankings and the evaluation for product champion in rubber industry is illustrated in Figure 2.



**C.R. = 0.06*

Figure 2. The evaluation for product champion selection

As a result from AHP model, passenger car tire is identified as the product champion of Thailand. Passenger car tire industry is one of the highest exporting values with high demand in global market. The following sections will depict a review of passenger car tire industry in Thailand and provide the direction for competitiveness improvement.

3. PASSENGER CAR TIRE INDUSTRY IN THAILAND

3.1 Current Situations

According to AHP model, passenger car tire industry was selected as the product champion of rubber industry in Thailand. The result accords with experts' opinion in the field of rubber industry. They consider that tire is one of the highest potential industries and should be supported.

A structure of passenger car tire industry supply chain is illustrated in Figure 3.

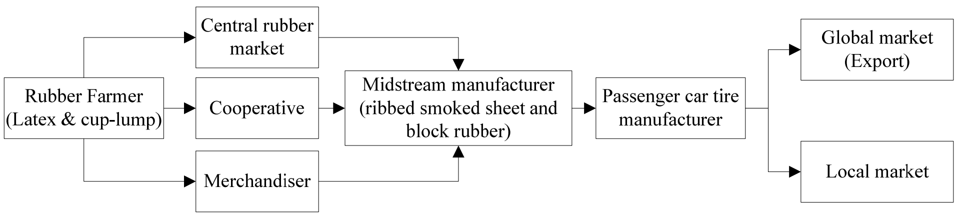
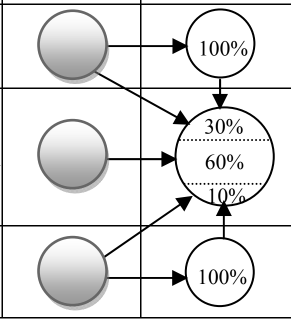


Figure 3. A structure of tire industry supply chain

In Thailand, rubber planting areas are located in Southern, Eastern, and Northeastern region. Mostly, rubbers are supplied to midstream firms nearby the planting areas. Then, natural rubber products from midstream manufacturing (such as ribbed smoked sheet and block rubber) are sent to local tire manufacturing. Passenger car tire manufacturers are mostly located in the Central and the East of Thailand.

From focus group and in-depth interview with midstream and downstream manufacturers in Thailand, the proportion of rubber products flow from upstream to midstream are shown in Table 3.

Table 3. The proportion of midstream products from raw material location

Location	Upstream	Midstream	Description
Northeastern region		100%	- Manufacturers in Northeastern region consume raw materials in their own area - Some of midstream products are sent to midstream plant in Eastern region
Eastern region		30% 60% 10%	- 60% of midstream products are produced by using raw material in their own area. - 10% and 30% of midstream products are produced by using raw material from Southern and Northeastern areas, respectively
Southern region		100%	- Manufacturers in Southern region consume raw materials in their own area - Some of midstream products are sent to midstream plant in Eastern region

Source: in-depth interview (2010)

From Table 3, mostly, midstream manufacturers use raw materials within their area. Except for Eastern region, this area consumes more raw materials from Northeastern and Southern areas. It can be seen that 30% of raw material fed into Eastern part are from Northeastern part. As a new growing area, there are still few midstream manufacturing plants in the Northeastern part.

Moreover, rubbers in Northeastern area are increasing as a result of the government's "The One Million Rai Rubber Plantation Project phase I" program in 2003 which can be harvested in 2010. We realise that the capacity of midstream plants in this area may not be adequate for support the rapid increase of rubber growth. This situation will affect the capacity balance and transportation cost in the passenger car tire supply chain in the near future.

4. THE DIRECTION FOR SUPPLY CHAIN IMPROVEMENT

4.1 Capacity Balance

From the problem stated above, we realize that to improve the passenger car tire supply chain, the product flows from upstream need to be managed. From the authors' previous work (Kritchanchai, 2009), it is found that there is no policy support and manage the product in this new growing area - Northeastern.

Most of rubber in Northeastern region flow to the Eastern part for midstream manufacturing. This results in over flown capacity at midstream plant together with high transportation cost. This urges us to consider about adding up capacity for midstream process in Northeastern region. That means number of midstream manufacturing must be increased in Northeastern region in order to consume its own area's raw material.

We then investigate the current situation capacity in each region in Table 4.

Table 4. The current situation capacity in each region

Location	Upstream dry rubber and cup-lump (tons)[1]	Midstream capacity available (tons)[2]	Midstream capacity status
Northeastern	104,601.43	13,980	Over capacity
Eastern	263,307.08	288,420	Slightly over capacity
Upper Southern	37,964.98	-	
Southern	2,330,378.25	3,162,936.00	Under capacity

[1] 92% of all upstream rubbers flow in the form of rubber latex and 8% is in the form of cup-lump.
90% of rubber latex supply to dry rubber industry and 10% supply to concentrated latex industry. (Department of Trade Negotiations, 2004)
[2] Capacity of Midstream plant (block rubber and ribbed smoked sheet) in 2006 (based on data from Department of Industrial Work (2009))

From Table 4, it can be seen that the midstream manufacturing in Northeastern is inadequate. Furthermore the rubber from the government project on "The One Million Rai Rubber Plantation Project" (1 rai=1,600 m^2) will be harvested in year 2016. Here we consider the production rate in year 2016 in comparison with the capacity available. The Northeastern production rate in year 2016 comes from 2 sources.

Firstly, we use trend analysis method to forecast the Northeastern outcomes in year 2016 which results in 327,910.90 tons. The forecasting of rubber production in year 2016 is shown in Table 5.

Table 5. The forecasting of upstream rubber production in year 2016

Region	P rovince	Equation	MAPE	Production 2016 (tons)	Region	Province E	quat ion	MAPE	Production 2016 (tons)
Northeastern	Loei	Yt = 7669.5 + 935.2*t	1	18,891.90	Eastern	Prachinburi	Yt = 66.0 + 165*t	11.69	2,046.00
	Nongbualumphu	Yt = 1076.5 + 241.5*t	3.31	3,974.50		Chachoengsao	Yt = 8697 + 2683*t	2	40,893.00
	UdonThani	Yt = 9624 + 1465.1*t	5	27,205.20		Srakao	Yt = 2230.0 + 15.8*t	1.18	2,419.60
	Nong Khai	Yt = 16931.5 + 4795.2*t	3	74,473.90		Chantaburi	Yt = 69707 + 4996*t	1	129,659.00
	Sakonnakhon	Yt = 2199.5 + 615.1*t	4.8	9,580.70		Trad	Yt = 40090 + 1704*t	2	60,538.00
	Nakhon Panom	Yt = 4619 + 930.9*t	3.2	15,789.80		Rayong	Yt = 109615 + 9306*t	1	221,287.00
	Mukdahan	Yt = 3627.5 + 859.1*t	4.8	13,936.70		Chonburi	Yt = 25348 + 2515*t	2	55,528.00
	Yasothon	Yt = 4619 + 930.9*t	3.2	15,789.80		**Tot**		**al**	**512,370.60**
	Amnatcharoen	Yt = -71 + 241.7*t	7.78	2,829.40	Southern	Chumporn	Yt = 102543 + 1181*t	2	116,715.00
	UbonRatchathani	Yt = 4541.5 + 1085.8*t	4	17,571.10		Ranong	Yt = 27846 - 1357.00*t	4	11,562.00
	Sisaket	Yt = 7444.5 + 1517.4*t	0.9	25,653.30		SuratThani	Yt = 446671 - 7989.50*t	4	350,797.00
	Surin	Yt = 2513 + 1925.5*t	12	25,619.00		PhangNga	Yt = 166861 - 6315.40*t	3	91,076.20
	Buriram	Yt = 5086 + 4091.2*t	8	54,180.40		Pukhet	Yt = 29413 - 1332.10*t	2	13,427.80
	Mahasarakam	Yt = 152.5 + 75*t	3.34	1,052.50		Krabi	Yt = 134529 + 893*t	0.2	145,245.00
	Roiet	Yt = 1760 + 154.9*t	1.89	3,618.80		Trang	Yt = 335984 - 1610.60*t	1	316,656.80
	Kalasin	Yt = 2538.5 + 586.8*t	1.65	9,580.10		Pattalung	Yt = 130099 + 3171*t	2	168,151.00
	Khonkean	Yt = 1360.5 + 129.1*t	2.41	2,909.70		Songkhla	Yt = 355711 + 3113*t	0	393,067.00
	Chaiyaphum	Yt = 923 + 154.8*t	4.68	2,780.60		Satun	Yt = 59902 + 1638*t	2	79,558.00
	Nakhonratchasima	Yt = 325.5 + 179*t	5.71	2,473.50		Pattani	Yt = 53539 + 3587*t	2	96,583.00
			Total	**327,910.90**		Yala	Yt = 201412 + 12085*t	2	346,432.00
						Narathiwat	Yt = 218471 + 9634*t	1	334,079.00
						Nakhon SiThammarat	Yt = 321634 - 9091.40*t	2	212,537.20
								Total 2,	**675,887.00**

Source: Final report “The Implementation of Logistics and Supply Chain Management in Thailand” (2010)

Secondly, the additional of 146,565.65 tons of rubber products from One Million Rai Project are included in the consideration (Rubber Research Institute of Thailand of Nong Khai province, 2010). Thus, the Northeastern production rate in year 2016 will be 474,476.55 tons. This again compares to the overall capacity shown in Table 6.

Table 6. The production rate in year 2016

Location	Upstream production in 2016 (tons)	Upstream dry rubber and cup-lump in 2016 (tons)[1]	Midstream capacity for dry rubber and cup-lump (if not increasing) (tons)
Northeastern	474,476.55 423,	100.23	13,980
Eastern	512,370.60 456,	891.11	288,420
Upper Southern	48,133.30 42,	921.43	
Southern	2,664,325.00 2,	375,831.89	3,162,936
Overall	3,699,305.45 3,	298,744.66	3,465,336

[1] 92% of all upstream rubbers flow in the form of rubber latex and 8% is in the form of cup-lump.
90% of rubber latex supply to dry rubber industry and 10% supply to concentrated latex industry.
(Department of Trade Negotiations, 2004)

It can be seen that when the outcomes from One Million Rai Project in new growing areas bloom, the midstream capacity for dry rubber and cup-lum p, which are the m ain raw m aterial for tire industry, will be shortage. This convinces the idea of adding up midstream plant in the new growing area - Northeastern region. The next question will be the suitable locations for the midstream manufacturing in the Northeastern region.

4.2 Location Allocation

Here, we determine the location of the cluster for allocate midstream manufacturing. It is expected that the investment and supports from the government should be encouraged.

Northeastern region is divided into group of provinces. In each group, collecting stations are determined. First, the break-even point distance for transportation from upstream to midstream is calculated based on the break-even point transportation cost for farmers in Northeastern area which is 5,299.33 THB/tons. Then, variable cost and fixed cost are considered. The distance of break-even point is resulted in 132 kilometers (at price of fuel 28 THB/litre) (Maersuntea and Wongsawat, 2006). After that, seven locations are selected as the location of midstream plants in this region namely; Nong Khai , Leoi , Udon Thani , M ukdahan, Kal asin, Si Saket , and B urirum provi nces. The l ocation of midstream plants in Northeastern region are illustrated in Figure 4.

Figure 4. The location of midstream plant in Northeastern region

This can be interpreted that the provinces within less than 132 kilometers should be grouped as a cluster. The quantity of rubber in the cluster is then consolidated at a representative province which has the largest amount of rubber in the group.

These seven locations should be able to absorb the production from upstream to midstream process. The recommended capacity in each location are calculated based on the rubber plantation area in each cluster. This is shown in Table 7.

Table 7. The recommended capacity of midstream plant in each location

Location	Production in year 2016 (tons)	Proportion (%)	Production from One Million Rai Project (tons)	Expected capacity in year 2016[1] (tons)
Nong Khai	66,409.87	22.71	29,683.08	96,092.95
Leoi	16,846.29	5.76	7,529.75	24,376.03
Udon Thani	30,398.20	10.40	13,587.02	43,985.22
Mukdahan	35,051.02	11.99	15,666.68	50,717.70
Kalasin	12,708.26	4.35	5,680.18	18,388.44
Si Saket	55,147.17	18.86	24,649.02	79,796.19
Burirum	75,843.91	25.94	33,899.79	109,743.70
Total	**292,404.71**	**100 130,**	**695.52**	**423,100.23**

[1] Capacity for block rubber and ribbed smoked sheet production in 2016

When the flows from Northeastern upstream have been redesigned to the midstream within the region while midstream capacity is provided, it is expected that the transportation cost will be reduced

The expected situation of the new flows is then compared to the current situation where midstream plants are located in Eastern region. It is found that the transportation cost will be decreased from 15,424,541.98 to 12,439,146.76 THB/year. This is calculated based on the proportion of 2.94% of total rubber used for tire industry downstream at Eastern and Central regions. Table 8 illustrates the cost saving of the new flows.

Table 8. The cost saving of the new flows

Scenario	Northeastern Products flow	Transportation cost (THB/kg)[1]		Overall Production flows (tons)[2]	Production flow for passenger car tire industry (tons)[3]	Transportation cost (THB)
		Upstream - Midstream	Midstream - Downstream			
1 (AS-IS)	Northeastern upstream – Eastern midstream – Eastern/Central downstream	0.9 0.	34	423,100.23	12,439.15 15,	424,541.98
2 (TO-BE)	Northeastern upstream – Northeastern midstream – Eastern/Central downstream	0.5 0.	5	423,100.23	12,439.15 12,	439,146.76

[1] In-depth interview, [2] Dry rubber and cup-lump in 2016

[3] Passenger car tire industry consume 2.94% of all production (dry rubber and cup-lump) in 2016 (based on data from Rubber Research Institute of Thailand (2009))

5. CONCLUSIONS

Passenger car tire industry is identified as one of the product champions in Thailand. It is one of the highest exporting values with high demand in global market.

In this paper, we concentrate on the flows of rubber products from upstream to midstream for tire industry. We found that the capacity of midstream plants in the Northeast of Thailand may not be adequate for supporting the rapid increase of rubber growth. This problem convinces that midstream plants in Northeastern area must be increased. For this industrial sector, strong developments and increasing the plant in this new growing area is a must.

Although, the current situation capacity could still absorb the upstream productivity, the outcomes from the government's project (The One Million Rai Project) will overflow the available capacity at midstream. The locations for midstream manufacturing plants in Northeastern region are recommended. It is expected that in this new growing area, connecting the upstream to midstream within the same region will result in reducing the supply chain cost.

For further research, the appropriate location and number of downstream plants will be studied. Then, the direction for supply chain competitiveness in this industrial sector for rubber products flow from midstream to downstream can be recommended.

ACKNOWLEDGEMENTS

The authors are grateful to the generous support from Thailand Research Fund (TRF).

6. REFERENCES

Bhushan, N. and Rai, K. (2004). *Strategic Decision Making Applying the Analytic Hierarchy Process*. Springer, United States of America.

Chauhan, K.A., Shah, N.C., and Rao, R.V. (2008). The Analytic Hierarchy Process as a Decision-Support System in the Housing Sector: A Case Study, *World Applied Sciences Journal, 3 (4):* 609-613.

Council of Logistics Management: Definition of Logistics. http://www.cscmp.org/

Department of Trade Negotiations: Rubber Industry. http://www.thaifta.com/

Friedland Global Capital News: World Tire Market Demand. http://friedlandglobalnews.com/

Kritchanchai, D., Somboonwiwat, T., Chaveesuk, R., Atthirawong, W., Choomrit, N., Wasusri, T., and Kingpadung, K. (2009). The evaluation of integrated industrial logistics system and supply chain management in Thailand. *Final Report, Thailand Research Fund (TRF),* Thailand.

Kritchanchai, D., Somboonwiwat, T., Chaveesuk, R., Atthirawong, W., Wasusri, T., and Supeekit, T. (2010). The Implementation of Logistics and Supply Chain Management in Thailand. *Final Report, Thailand Research Fund (TRF),* Thailand.

Kritchanchi, D. (2009). Rubber Supply Chain in North Eastern Part of Thailand. *Proceedings of Asia Pacific Industrial Engineering & Management Systems Conference 2009, Kitakyushu, Japan,* 575-581.

Maersuntear, J. and Wongsawat, P. (2006). The study of potentiality for rubber industry along the East-West economic corridor in the Northeast of Thailand. Department of Industrial Engineering, Khon Kaen University.

Michelin: Factbook. http://www.michelin.com/

Office of National Economic and Social Development Board: Industry Profile Report. http://www.nesdb.go.th/

Rubber Research Institute of Thailand: Thailand rubber statistics. http://www.rubberthai.com/statistic/stat_index.htm

The Office of Industrial Economics: Industry Statistics. http://www.oie.go.th/industrystat_th.asp

Tseng, Y.Y., Yue, W.L., and Taylor, M.P. (2005). The Role of Transportation in Logistics Chain. *Proceeding of The Eastern Asia Society for Transportation Studies*, *5:* 1657-1672.

UN Comtrade: United Nations. http://comtrade.un.org/db/

A SIMULATION DESIGN METHOD OF REFRIGERATED WAREHOUSES USING AN ASPECT-ORIENTED APPROACH MODELING*

Gyu-sung Cho[1*], Young-jun Lee[2], Dae-wook Lee[2], and Ho-gyun Kim[2]

[1]Department of Logistics System
Tongmyong University
179, Sinseonno, Nam-gu
Busan, Korea 608-711
e-mail: gscho@tu.ac.kr

[2]Department of Industrial & Management engineering
Dongeui University
995 Eomgwangno busnajin-gu
Busan, Korea, 614-714
e-mail: hgkim@deu.ac.kr

Abstract: Refrigerated warehouses play a buffer function in logistics system to meet the various demands of consumers. Over the 50% of Korean refrigerated warehouses are located in Busan and Busan has become a strategic region for cold chain industries. This paper suggests an aspect-oriented approach to simulation design and analysis in system design for the refrigerated warehouses. The purpose of this paper is to formulate a new methodology for developing reusable, extendible, and modifiable control software for a refrigerated warehouse model using aspect-oriented approach Modeling.

1. INTRODUCTION

An aspect-oriented approach modeling (AOAM) consists of a primary model and one or more aspect models. An aspect-oriented approach modeling has a feature that crosscuts the primary model. AOAMs are generic descriptions of crosscutting features that must be instantiated before they can be composed with the primary model based on the metamodel for aspect- oriented modeling[Chavez et al., 2001, 2002, Kim et al., 2007]. The term of the AOAM can be used to demonstrate the space of programmatic mechanisms for expressing crosscutting concerns[Kiczales et al., 1997]. AOAM should be built upon a conceptual framework and is to denote the space of modeling elements for specifying crosscutting concerns at a higher level of abstraction. In this paper, we illustrate how a signature-based composition approach can be used to compose class models and describe how composition directives can be used to ensure that the composition approach produces desired results. We have developed a prototype system which implements the class model composition behavior specified in the composition metamodel to design the refrigerated warehouse system. This paper suggests a simulation design and analysis in system design for the refrigerated warehouses using the AOAM. The purpose of this paper is to formulate a new methodology for developing reusable, extendible, and modifiable control software for a refrigerated warehouse model.

2. DEFINING THE ASPECT-ORIENTED APPROACH MODELING

In this paper, we demonstrate the mixed concept with object-oriented model and aspect-oriented model. Usually the object- oriented model defines a data object as containing code and data. Generally, code and data have been kept apart. In an object-oriented data model, the code and data are merged into a single indivisible thing which is an object. Figure 1 shows the backbone suggested for structural modeling. Figure 2 is the basic flow chart based on the AOAM.

Figure 1. Concept of the AOAM

Figure 2. AOAM Flow Chart

"Base Elements" describes modeling elements that may participate in a "Crosscutting" relationship with one or more elements of type "Crosscutting Element". "Crosscutting Element" describes modeling elements that may participate in a "Crosscutting" relationship with one or more elements of type "Base Element". "Crosscutting" is a relationship that denotes a new extension mechanism for combining incremental descriptions of structure and behavior localized in "Crosscutting Element" into a "Base Element". "Aspects" are defined as parameterized model elements. Aspects comprise a set of local features and a set of crosscutting interfaces. A "crosscutting interface" models an archetypical representation of a set of crosscut objects under the aspect's perspective. A "crosscutting feature" is a model element that describes a feature to be combined to one or more base elements. Aspect operations which refine or override are specified as parameterized elements, and they contain one or more unbound parameters which represent operation name and signature.

3. REFRIGERATED WAREHOUSES OPERATIONS

Refrigerated warehouses are facilities where perishable foodstuffs are handled and stored under controlled temperatures with the aim of maintaining quality. Refrigerated warehouses play an essential role in the food delivery logistics system.

It i s ve ry i mportant t o c onsider t he operations o f refrigerated st orages si nce t he c onsumptions of fresh foods a re increasing.

Storing process into a refrigerated warehouse is defined as follows:
(1) Freights arrival at the warehouse via trailer or truck.
(2) Pulling the trailer or truck into the dock and after confirming attached "Customer Seal and Shipping Seal"
(3) Loading the Freights on the pallet.
(4) Attaching the tags with the product information such as storing place, quantity and date.
(5) Using a forklift to move the pallet to elevator or vertical lift in the warehouse.
(6) Transferring them into the warehouse.

Releasing process from a refrigerated warehouse is defined as follows:
(1) Receiving the requests from shippers.
(2) Searching for the freights by using the information of tags.
(3) Looking for the pallets where the freights are loaded.
(4) Transferring the requested pallets to the gates and loading the pallets into trucks.

4. REFRIGERATED WAREHOUSES SYSTEM DESIGN AND REALIZATION

In this paper, we apply the concept into a case study which evaluates the system of a re frigerated warehouse located in Busan. Fi gure 3 sh ows t he i ntegrated design m ethod using o bject-oriented m odeling a nd as pect-oriented m odeling t o show t he system . Pu rposed a syste m can be b uilt b y m ixing object-oriented m odeling an d aspect-oriented ap proach modeling and design the system to realize an accurate analysis.

We develop a simulation model using AutoMod simulator considering the AOAM for the performance analysis of a ref rigerated warehouse system. W e i mplemented t he s imulation m odel usi ng t he following design pa rameters as follow:

System configurations of a refrigerated warehouses system are defined as follows:
- Number of bin in the warehouse = 30*6 bin = 180 bin
- Total stored pallets = 180 * 21 = 3,780(supposition: stored 21 pallets within each bin)
- Each bin sizes = 30m*1m*3m
- Forklift speed = 50m/min
- Loading & unloading time per one pallet = 30sec
- Arrival time of freights = exp (30) sec
- Order time of freights = exp (5) min
- Capacity of one forklift = 1 Pallet
- Size of warehouse = 30m*52m
- Hallway size = 5m*52m

In this m odel, the system can be flowed such as figure 4. Freights and forklifts which are depicted in figure 4 are always research ing th e workload and work lists fro m th e aspect-oriented m odeling and fu lfill th e ord ers by obj ect-oriented modeling.

Method	Conveyor		Forklift		Rack	
Object-oriented modeling	Section Type Width Accumulation Motor Velocity Acceleration Deceleration	Station Type Capacity Align Protoeye Type Block timeout value Cleared timeout value	Guide Path Guide Path Transfer Control Point	Vehicle Capacity Pick up time Set down time Number Acceleration Deceleration Forward velocity Reverse velocity	Rack Rack Number of aisle Aisle width Number of bay Bay width Bay depth Next tier height P&D stand Zone	SRM Capacity Pick up time Set down time Number H_acceleration H_deceleration V_acceleration V_deceleration
Aspect-oriented modeling	Process P_setting() P_entering() P_coming() Queue Q_input Q_output Load L_Loadtype L_Loaditems Variable V_input_to_conv() V_output_to_conv() Random Number Streams Run control		Process P_setting() P_entering() P_coming() Queue Q_input Q_output Load L_Loadtype L_Loaditems Variable V_input_to_Forklift() V_output_to_Forklift() Random Number Streams Run control		Process P_setting() P_rack_entering() P_rack_outting() Queue Q_rack_input Q_rack_output Load L_loadtype L_Loaditems Variable V_rack_to_PD() V_rack_to_zone() Random Number Streams Run control	

Figure 3. Integrated Design Method using Object-Oriented Modeling and Aspect-Oriented Modeling

Figure 4. Operating Flow Chart of a Refrigerated Warehouses

Figure 5 and 6 show the order picking warehouse system layout demonstrated by AutoMod simulator.

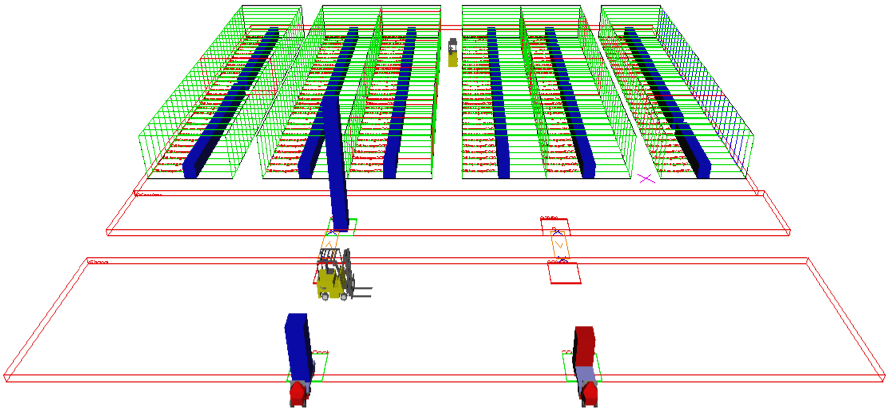


Figure 5. Warehouse System Layout

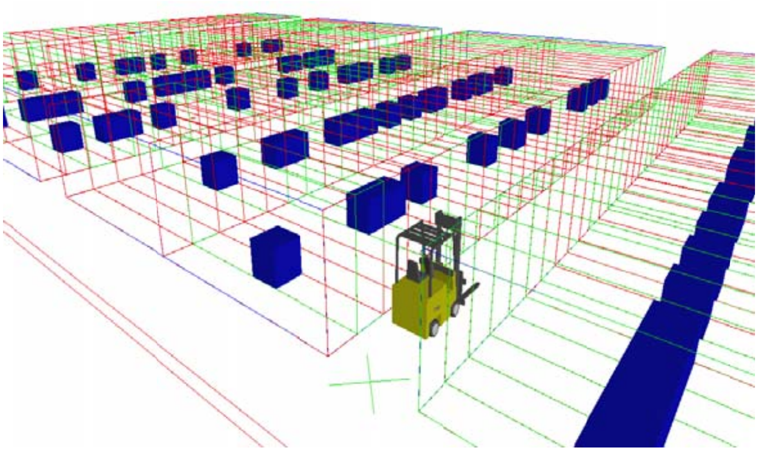


Figure 6. Operating the Forklifts in a Warehouse

We obtained the results by using the AOAM. The results are shown in the table 1 and 2. At the gates of the warehouses, the average rates of delivering/retrieving the freights by forklifts is about 42.1% and parking rate is 57.9% . In the average rates of delivering/retrieving the freights by forklifts is about 100% and parking rate is 0% . The capacities of forklifts are 15.3 pallets per 1 hour. Total handling quantizes are 5,109 pallets and the average share of racks is 44% and the average share of a rack is 9.18 pallets. More detailed results such as storage utilizations, minimum, maximum, average levels and total handling quantities are showed in appendix.

Table 1. Performance of the Forklifts

Division	Delivering			Retrieving			Parking
	Freights handling capacity (/hr)	Unit operation time (/min)	Rate	Freight handling capacity (/hr)	Unit operation time (/min)	Rate	Rate
Gates of the warehouse	15.3 0	.65 1	6.5% 1	5.3	1 25	.5% 5	7.9%
Within the warehouse	15.22 3	.11 7	8.8% 1	5.22 0	.75 1	9%	0

Table 2. Performance of the Racks

Division Sm	allest	Maximum	Average	Total	Average Util.
Storage	0.99	16.28 9.	18 5,	109	44%

5. CONCLUSIONS

In this paper, we developed a simulation model using AutoMod simulator considering the AOAM for the performance analysis of refrigerated warehouse system. The purpose of this paper is to formulate a new methodology for developing reusable, extendible, and modifiable control software for a refrigerated warehouse model. For that, this study shows an aspect-oriented structure to a nalyze th e p erformance o perated in th e warehouse and a simulation m odel is showed to demonstrate t he co ncept. S uch a n e xtracted aspect m odules can be a pplied i nto t he st ructure o f t he existing sy stem without change or modification because it is independence.

ACKNOWLEDGEMENTS
This material is based upon work partially funded by Young Nam Sea Grant.

6. REFERENCES

Brooks Automation. (2005). *AutoMod 12.0 User's Guide*, USA.

Chavez, C. and Lucena, C. (2001). Design-level Support for Aspect-oriented Software Development, *In International Workshop on Advanced Separation of Concerns at OOPSLA.*

Chavez, C. and Lucena, C. (2002). A Metamodel for Aspect-Oriented Modeling, 1-6.

Kiczlaes, G., Lamping, J., Mendhekar, A., Maeda, C., Lopes, C., Loingtier, J. an d Irwin, J. (1997). Aspect-Oriented Programming, *European Conference on Object-Oriented Programming*, 1241: 220-242.

Kim, T., Um, I. and Lee, H. (2007). The Simulation Design and Analysis Method of Integrated Logistics System using an Aspect Oriented Approach, *IE Interfaces*, 20(4): 438-447.

Appendix:

Table 3. Performance of the each rack

Storage A	Min.	Max.	Ave.	Total	Util.	Storage B	Min.	Max.	Ave.	Total	Util.
1	0	11	4.10	30	0.195	1	0	8	3.12	27	0.148
2	0	18	7.35	32	0.350	2	0	14	5.74	26	0.273
3	0	10	3.12	27	0.148	3	0	9	4.44	30	0.212
4	0	7	2.60	22	0.124	4	0	11	5.03	21	0.239
5	0	13	7	23	0.333	5	0	16	7.09	31	0.337
6	0	16	8.89	30	0.423	6	0	21	9.92	31	0.472
7	0	14	6.90	30	0.329	7	0	15	6.64	30	0.316
8	0	10	3.79	23	0.181	8	0	14	6.79	28	0.323
9	0	15	6.75	27	0.321	9	0	16	6.95	28	0.331
10	0	15	7.91	24	0.377	10	0	19	9.33	28	0.444
11	0	20	10.20	25	0.486	11	1	19	12.53	27	0.597
12	1	18	8.93	26	0.425	12	2	21	13.26	31	0.631
13	0	17	8.33	23	0.397	13	0	21	14.23	26	0.678
14	1	18	10.83	24	0.516	14	1	14	9.28	21	0.442
15	2	20	9.96	36	0.474	15	0	12	6.70	34	0.319
16	0	21	12.69	25	0.604	16	0	21	11.95	24	0.569
17	0	16	9.33	25	0.444	17	1	16	9.30	34	0.443
18	0	9	3.63	21	0.173	18	0	8	3.56	26	0.169
19	2	21	9.61	29	0.457	19	2	21	12.68	31	0.604
20	0	15	6.30	37	0.300	20	0	9	4.70	37	0.224
21	1	21	9.35	37	0.445	21	0	13	5.37	21	0.256
22	0	15	7.21	18	0.344	22	0	18	8.53	27	0.406
23	1	14	7.82	26	0.372	23	2	21	11.40	31	0.543
24	0	18	9.91	38	0.472	24	0	13	7.94	32	0.378
25	0	7	4.27	20	0.203	25	0	10	4.39	22	0.209
26	0	14	6.78	21	0.323	26	0	16	9.50	25	0.452
27	2	21	15.30	27	0.729	27	2	13	7.05	16	0.336
28	0	15	6.63	33	0.316	28	1	12	6.21	37	0.296
29	1	17	9.36	30	0.446	29	1	13	9.93	24	0.473
30	0	15	8.62	22	0.411	30	1	21	12.53	29	0.597

Storage C	Min. M	ax .	Ave.	Total	Util.	Storage D	Min. M	ax .	Ave.	Total	Util.
1	0	18 1	1.74	30 0	.559	1	1	21 1	2.04	32 0	.573
2	0	20	9.12	29 0.	434	2	0	17 10	.10	23 0.	481
3	1	21 10	.07	34 0.	480	3	2	13	7.61	29 0.	363
4	1	16	6.40	25 0.	305	4	0	21 14	.77	33 0.	703
5	1	16 7.	94	23	0.378	5	1	13 7.	53	22	0.358
6	3	7	5.00 18		0.238	6	3	15	8.34 28		0.397
7	1	13 7.	25	28	0.345	7	2	20 9.	26	27	0.441
8	0	11 4.	67	36	0.222	8	0	13 6.	48	34	0.308
9	1	8	5.89	20 0	.281	9	1	18 1	0.66	29 0	.507
10	2	21 15	.37	31 0.	732 1	0	0	21 13	.59	32 0.	647
11	0	21 13	.59	25 0.	647 1	1	0	21 10	.65	28 0.	507
12	0	21 11	.13	28 0.	530 1	2	1	21 13	.40	28 0.	638
13	1	18 11	.55	21 0.	550 1	3	1	16 10	.33	20 0.	492
14	2	19 11	.48	28 0.	547 1	4	1	17	8.34	24 0.	397
15	1	18 11	.10	27 0.	528 1	5	2	21 14	.60	35 0.	695
16	6	17 13	.92	27 0.	663 1	6	1	21 14	.40	24 0.	686
17	0	19	7.55	30 0.	360 1	7	1	14	7.79	31 0.	371
18	0	16	8.54	29 0.	407 1	8	0	12	6.17	22 0.	294
19	0	17 10	.51	20 0.	501 1	9	1	16 10	.43	22 0.	496
20	2	14	8.34	28 0.	397 2	0	5	21 13	.46	48 0.	641
21	1	11	5.71	31 0.	272 2	1	0	12	7.46	30 0.	355
22	2	19 11	.45	35 0.	545 2	2	3	18	9.85	30 0.	469
23	2	15	8.13	28 0.	387 2	3	3	21 15	.02	36 0.	715
24	0	11	5.32	36 0.	253 2	4	0	9	2.58	30 0.	123
25	0	21 11	.10	32 0.	528 2	5	1	15	8.87	26 0.	422
26	5	14 10	.12	36 0.	482 2	6	1	20 10	.99	33 0.	523
27	1	20 12	.60	28 0.	600 2	7	3	16 11	.25	23 0.	536
28	3	14	8.55	39 0.	407 2	8	0	12	5.56	37 0.	265
29	1	15	6.84	29 0.	326 2	9	1	16	9.16	30 0.	436
30	1	21 11	.15	28 0.	531 3	0	3	21 12	.21	26 0.	581

Storage E	Min. M	ax .	Ave.	Total	Util.	Storage F	Min. M	ax .	Ave.	Total	Util.
1	3	21 1	2.27	32 0	.584	1	2	18 1	0.18	29 0	.485
2	1	17 1	1.11	21 0	.529	2	1	21 1	2.75	29 0	.607
3	4	14 1	0.06	29 0	.479	3	2	20 1	2.06	32 0	.574
4	0	14 6.	63	24	0.316	4	0	14 6.	82	25	0.325
5	1	12 4.	83	25	0.230	5	1	14 8.	69	26	0.414
6	1	19 1	2.99	35 0	.619	6	1	21 1	1.97	32 0	.570
7	0	8	3.96 24		0.189	7	0	13	5.85 29		0.279
8	0	13 5.	95	36	0.283	8	0	12 5.	33	34	0.254
9	4	21 1	4.09	33 0	.671	9	3	19 1	2.68	25 0	.604
10	0	16	6.97	26 0.	332 10		1	16	8.94	27 0.	426
11	1	17	9.54	23 0.	454 11		0	21 14	.25	29 0.	679
12	0	12	7.23	29 0.	344 12		0	21	9.02	30 0.	429
13	1	16 11	.89	26 0.	566 13		5	21 11	.55	32 0.	550
14	1	20 12	.79	25 0.	609 14		6	18 12	.22	30 0.	582
15	5	18	9.17	35 0.	437 15		2	21 11	.73	33 0.	558
16	2	20 12	.88	25 0.	613 16		1	21 13	.90	25 0.	662
17	0	12	5.86	26 0.	279 17		2	12	7.03	28 0.	335
18	0	19 12	.17	43 0.	579 18		1	14	8.74	30 0.	416
19	3	19 11	.11	29 0.	529 19		1	21 12	.85	26 0.	612
20	0	15	6.53	32 0.	311 20		0	14	6.61	37 0.	315
21	1	21 14	.22	24 0.	677 21		2	21 12	.84	29 0.	611
22	2	17 11	.33	29 0.	540 22		2	18	9.10	28 0.	433
23	1	15	7.40	24 0.	352 23		1	13	7.75	26 0.	369
24	0	21 12	.53	26 0.	597 24		0	18 10	.36	30 0.	493
25	1	14	8.82	30 0.	420 25		2	16	9.68	31 0.	461
26	4	21 14	.52	38 0.	691 26		0	18 10	.62	29 0.	506
27	0	15	9.44	19 0.	450 27		1	20 11	.62	28 0.	553
28	1	17	7.18	32 0.	342 28		3	16 10	.32	32 0.	491
29	1	21 11	.58	38 0.	552 29		0	15	9.03	27 0.	430
30	0	19	9.68	27 0.	461 30		0	16	7.89	19 0.	376

SECURITY EVENT MANAGEMENT

Rainer Müller[1]

[1]Institute of Shipping Economics and Logistics (ISL)
t.i.m.e.Port II
Barkhausenstraße 2
Bremerhaven, Germany
e-mail: mueller@isl.org

Abstract: This paper describes a Security Event Management system which adapts the SCEM (Supply Chain Event Management) concept and Fuzzy Logic on the field of container security, in order to improve the security of container transports.

After an introduction of security in container transports the different kinds of security related static criteria and dynamic events which could occur during container transports are explained. Furthermore, the concept of the software is explained.

The software system has been developed during a research project by the Institute of Shipping Economics and Logistics (ISL) and takes care of all security related criteria and events in the intermodal chain. The system registers the transport information of a container. During the transport, the system receives events which will be used for computing a security risk value for each container. By adaption of the Fuzzy Logic concept this value will be translated into a qualitative linguistic result.

1. INTRODUCTION

As far as security is concerned, the optimisation of container transportation relies primarily on a system of trust, where sealed container move unimpeded through the supply chain, noted in (APL Logistics 2003). But during a transport, different types of threats can occur, e.g. a seal of a container may be broken or a container may be transported to an unforeseen location. This implies the risk of theft of goods or usage of the container for acts of terrorism.

This paper describes a project of ISL developing a demonstrator to support transport organisers to help increasing the security of supply chains of container transport, as well as to decrease the risk of thefts or usage of the container for acts of terrorism.

First the Supply Chain Event Management (SCEM) approach and basics of Fuzzy Logic is explained. After the description of security related criteria and events, the concept of the software demonstrator and its implementation of SCEM and Fuzzy Logic is explained.

After introducing several ways how to generate the security related events, the calculation of a security value will be described.

2. THE SCEM APPROACH

The Supply Chain Event Management (SCEM) methodology for planning and monitoring supply chains is well known in the production industry. For the adoption of this concept to intermodal container transports the so-called events, which occur during a transport, have to be divided into two categories:

- Expected events such as loading and unloading messages. Each of those refers to an expected milestone of the transport process.
- Unexpected events, which may allude to problems such as delay messages or a notice indicating a problem.

While traditional systems for planning, tracking and tracing are passive, i.e. the user gets updated information only if he looks into the different systems; the idea of SCEM systems is to play an active role. To this end, the planned logistics chains (e.g. in door-to-door transports) with expected events at expected times will be defined and matched against incoming event messages. Depending on the configuration, certain triggers will be fired:

- either if expected events do not happen
- or other events occur which were not expected.

SCEM systems compare the expected milestones of the transport with the occurred events and decide on appropriate actions, e.g. inform the chain manager proactively. So, the user may stay passive, because the system plays the active role and he is informed only in case of problems. These problems can be coped with at an early stage, because the chain manager is informed after the occurrence of these problems. This enables the chain manager to reduce the impact of the problem on the transport chain.

It is important for an SCEM system to obtain the events automatically and timely, e.g. via EDI (e.g. discharge messages from port community systems), via mobile devices (e.g. handheld computers using GPRS communication) or using web interfaces.

3. FUZZY LOGIC

Fuzzy Logic enables software systems to deal with values which are not crisp. In detail, these values are called Fuzzy set, to be more precise, a Fuzzy set is a collection of elements where every element has a degree of membership. In detail, the assignment of an element to a fuzzy set describes a grade of membership ranging between zero and one, which is also defined as membership function (Zadeh, 1965). By contrast, in classical sets an element either belongs or does not belong to a set. In detail, the value of a membership function of classical sets is described by values of zero or one.

Fuzzy sets are extended classical sets, because the notions of inclusion, union, intersection, complement etc. are also applicable upon fuzzy sets, whereas fuzzy sets permit the ability to deal with imprecise values.

Figure 1. Fuzzification of value 20 %

Figure 1 shows the usage of Fuzzy sets in the field of risks. In detail, a crisp value of 20 % is translated into a Fuzzy set using the membership function. As a result the set belongs with a factor of 0.2 to negligible and with a factor of 0.8 to low risk.

4. SECURITY EVENT MANAGER

4.1 Introduction

In the following chapter the concept of a software system is described which monitors container transports in order to compute a security risk value. By adaption of the SCEM methodology the system is able to proactively inform the user about threats and security related problems. Furthermore, this system uses Fuzzy Logic which enables the software to deal with imprecise values.

For the computation of a security value the system has to know static information about the transport and security related events. At first, a list of a priori defined security related data and events will be presented. Afterwards, results of a workshop are described which show a list of security related criteria and events defined by the experts. Furthermore, different ways of obtaining security related events are described. The chapter closes with a section on computation of the security value.

4.1 Security related data and events

For a security system we have to distinguish between static information and events which occur dynamically.
The following priori-defined list of static criteria describes which kind of data is interesting for a security system:

- Value of goods: The higher the value the higher is the risk of theft.
- Type of cargo: dangerous goods could be used for acts of terrorism.
- Destination: There are locations which are more interesting for terrorism acts like others, e.g.: terminals.

- Companies: The kind of involved companies during the transportation has a great impact on the security, e.g. if they are screened and certified. The more companies are involved in an intermodal transportation the higher the risk of theft or terrorism acts.

Beside these static criteria there are dynamic events occurring during a transport which have an impact on the security of the container transport, for example:

- The truck with the container has left the assigned route
- Standstill of the truck with container on an unsecure parking lot
- The container has been opened unauthorized

Security related events can be messaged by different systems and persons.

<u>4.1.1 Automatic messaging of events</u>

By using smart units or smart containers events could be messaged automatically. These devices are able to detect breaking the door by light and contact sensors. Furthermore by using shock, fire, smoke and temperature sensors they can notify changes inside the container which could be a disruption. Some units are able to send a message by GSM/GPRS to a security system.

Some of those devices are also using RFID technology to track the containers position at important locations. There are also smart units and smart containers which are equipped with a GPS Module which could be used for permanent position tracking.

Due to the high initial and maintenance costs of these devices they seem to be useful only for containers with high-value goods or dangerous cargo.

Instead of such expensive units which are applied on each container, ISL has developed a low-price solution which can be used on the vehicle, like train, truck or barge. A standard mobile phone and a GPS receiver (or in the future Galileo) is used to track the container while it is transported by the equipped carrier. A small Java application on the mobile phone is receiving the GPS position by the GPS receiver and sends it by using web services to a server. Users are able to track the current position in a Google Maps application.

The current position of the container can be used for generating events, for example when a transported container enter or leaves a defined area, e.g. a terminal. The user is able to define the position and the radius of locations (like terminals or also motorway junctions) as well as the type of event which should be fired by entering or leaving the defined circle.

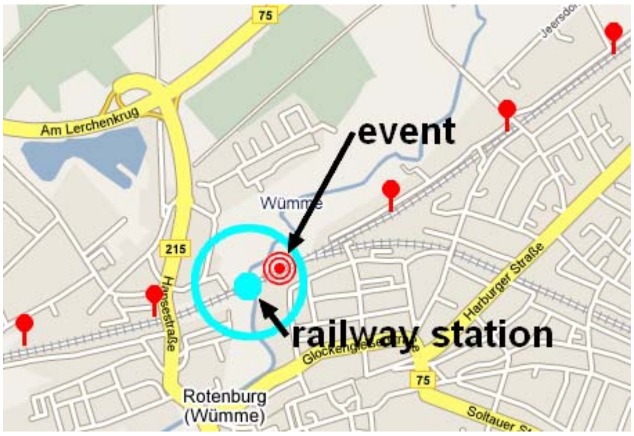


Figure 2. Generating events by entering an area

A server is receiving the GPS position and calculating if the transported container has entered or left a defined location and fires the respective event (cp. figure 1).

The user could assign a detailed description of the container´s route, which will be used as a security corridor. By comparing the current position and the security corridor the system will be able to identify a deviation and to create an event if necessary, e.g. a truck is leaving the motorway before it was planned (cp. figure 2).

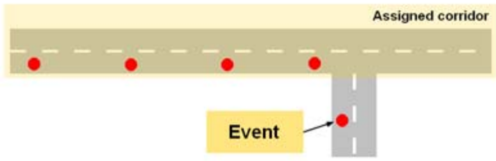


Figure 3. Detecting events by GPS

Furthermore, the system will be able to recognize standstills when the carrier is not changing its position for a longer time. There are several restrictions concerning such standstills:

- A standstill somewhere in a forest is more dangerous than on a motorway service area
- The longer the time the higher is the chance of a disruption
- A standstill during the night could be more dangerous than in the daytime

The system also has to take care of events which have occurred in the past, because e.g. the chance of a disruption is high if the container has left the assigned security corridor one hour ago and it is not moving anymore.

4.1.2 Generating events manually

It is handy to generate events in an automatic manner, but there are events which could only be detected by humans. Furthermore not every container will be equipped with a smart unit.

By using software and hardware clients, especially mobile clients, employees could send messages to the Security Event Manager which has to evaluate if an event should be created. Security related events could be:

- Changed goods, which could be detected by opening the container during a customs check
- False, missing or broken seal
- The carrier, chassis or driver is a different one than the planned one
- High security level of a terminal (in cause of ISPS)

4.3 Experiences of experts

In order to verify the a priori-defined list of security related criteria and events a workshop has been conducted where the participants like policemen, transport organisers and other security experts exchanged their experiences in container security and defined a list of security related criteria and events for the risk of theft. Furthermore, a questionnaire has been carried out in order to identify the impacts of these criteria and events on the security of current transport chains.

The impact describes the influence of a criterion or an event on the security of the transport. Following the Fuzzy Logic concept the values for the impacts on the security of the container transports are linguistic terms:

- negligible
- low
- medium
- high
- extreme

In the following security related criteria and events have been identified by the experts. Furthermore, the impact of the criteria or event on the security value is described.

4.3.1 Value of goods

The higher the value the higher is the risk of theft.
This static criterion has a high impact on the security.

4.3.2 Disposability and desirability of goods

Desirability describes how many persons are interested in buying that kind of stolen good and disposability says if a stolen good is easy to sell or not. These two values are referring to each other and have a high impact.

In most cases the disposability and desirability of a good is high if the value is also high (e.g. mobile phones), but for example a machine for special purposes could be of high value but it could be very hard to sell.
Disposability and desirability has the highest impact on the security of a container transport.

4.3.3 Involved parties

The kind of involved parties has a high impact on the security. If the parties are Authorised Economic Operators (AEO) the risk of theft is lower.

4.3.4 Information security

Information security has a high impact on the security. The experts stated that for example the truck driver should not know any details of the transport.

4.3.5 Physical security

Container transport itself is more secure than other transports. Usage of seals and eSeals could improve the security. If an eSeal is broken an event could be messaged which has a medium to high impact on the security.

4.3.6 Modal change

In intermodal transports containers are especially endangered at a modal change which has a medium impact.

4.3.7 Routing / Timeline

The route could be criterion with a medium impact on the transport, because there are regions which are more endangered for the risk of theft than other ones.

Events like standstill events or if a container is delayed have also a medium impact on the security of a chain.

4.3.8 Priori-defined list and experiences of experts

Most of the a priori-defined criteria and events have been confirmed by the experts.

4.4 Security value

The Security Event Manager has to calculate a security value assessing the current risk level of a transport to help the transport organisers to decide if the transport is endangered or not.

The security risk value cannot be a sharp precise value because on the one hand there is no sharp threshold value indicating the risk level. On the other hand there are no precise rules defining which event will lead to a certain risk level. So the algorithm which computes the security value has to be able to deal with experiences of experts which are described in linguistic terms but not in mathematically precise functions.

The system will use Fuzzy Logic to compute a qualitative linguistic result, e.g. “low risk” or “high risk”. This result will be computed by a function with two parameters:

- The static impact of a criterion, which describes the importance of a criterion, e.g. the importance of the value of goods, for risk of theft
- The dynamic value of the occurred event, e.g. the current position of the container is outside the assigned corridor

Figure 4a. Impact of a criterion

Figure 4b. Value of 20% is fuzzificated

Figure 4c. Calculating centre of gravity of each fuzzy set and calculating result with mean of maximum function.

Figure 4d. Fuzzification of result

Figure 4 illustrates an example of the calculation. At first, the impact of a criterion, for example "Disposability and desirability of goods" is translated into a fuzzy set (cp. Figure 4a). And second the disposability value of 20 % for the given kind of good will be fuzzificated; which means it is translated into a fuzzy set (cp. Figure 4b). Next, using Fuzzy Logic algorithms, the centre of gravity of each fuzzy set is calculated. Furthermore, the mean of maximum of both fuzzy sets is calculated, leading to a crisp result of 50 (cp. Figure 4c). In the end, this crisp result is fuzzificated into a fuzzy set in order to display the user the qualitative linguistic result "medium" (cp. Figure 4d). To sum up, the result of "Disposability and desirability of goods" for the current kind of good is "medium".

The calculation for each criterion and event creates one single result. In the future, it is planned to calculate an overall result which indicates the user an overall security value.

The functions of the Fuzzy Logic algorithm which determine the result for the events and their restrictions will be defined with the help of transport organisers and security experts. Their experiences and opinions will be translated into linguistic terms, e.g. for the influence of the value of goods concerning the risk of theft could be "high impact". These linguistic terms will be used by the Fuzzy Logic algorithm to compute the security value.

During operation, the staff of the transport organisers will be able to supply their own experiences to improve the efficiency of the system.

By now, ISL plans to define more than one security value. There will be one value for the risk of theft and a different one for the risk of acts of terrorism, because the risks are different. On the one hand e.g. the influence of the value of goods for the risk of theft is very high, but on the other hand for a terrorist this value is unimportant.

The user will have the option to declare a container transport to be endangered itself, as a result of circumstances, e.g. a very important container. This would change the threshold in the algorithm which will increase the security value accordingly.

Following the SCEM idea the security factors themselves will be used for generating messages, e.g. Mail or EDI messages, to proactively inform the user. The user himself will be able to define the content of these messages, which could contain simple information and also appropriate instructions in case of critical situations.

4.5 Security Event Manager Demonstrator

ISL has implemented the first prototype of the Security Event Manager, which receives planned data and events from different test systems and is able to compute a security risk value using Fuzzy Logic. A screenshot of the graphical user interface of the prototype is shown in figure 5.

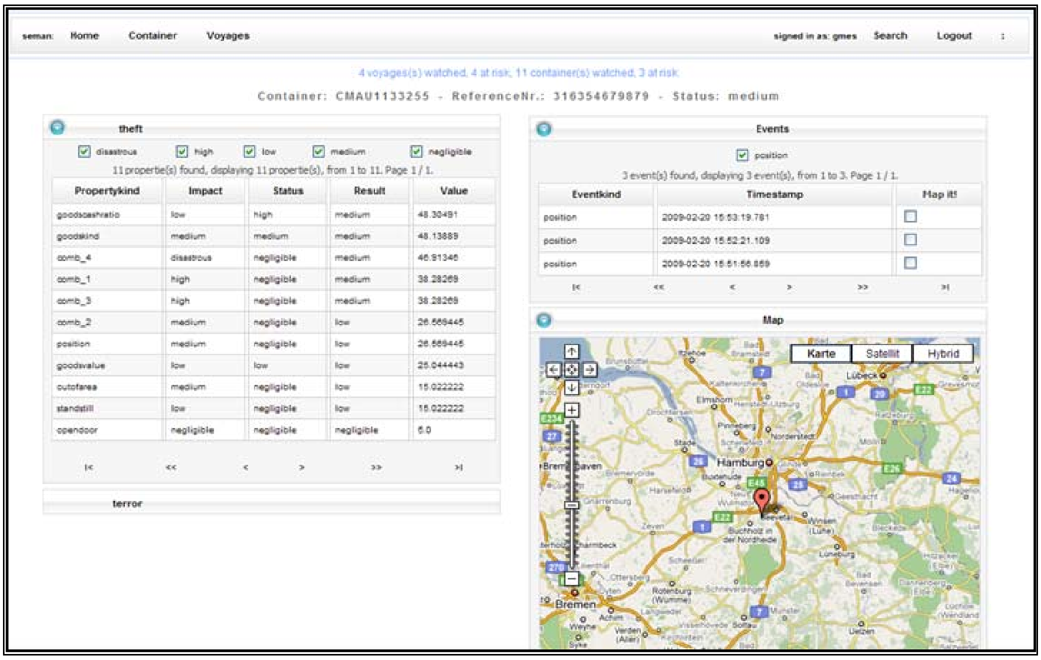


Figure 5. Screenshot of Security Event Manager demonstrator

The development process is still ongoing. In the next steps the user interface and the calculation of the security value will be improved.

5. CONCLUSION

It was shown that a Security Event Management system is able to compute a security risk value which can help to increase the security of intermodal container transport. By adapting the SCEM methodology the system can proactively inform the user which enables transport organisers to cope with problems at an early stage.

Instead of dealing with crisp values the Security Event Manager system is able to deal with experiences of experts by using qualitative linguistic terms which are computed by a fuzzy algorithm.

6. REFERENCES

APL Logistics (2003) Adding Security and Value to the Supply chain, http://www.apl.com/news/documents/security_white_paper.pdf

A.T. Kearney (2005): Smart boxes – RFID can Improve Efficiency, Visibility and Security in the Global Supply Chain, Chicago 2005

Collins, J. (2005) IBM, Maersk Developing Cargo Tracker, RFID Journal Sept. 22, 2005, http://www.rfidjournal.com/article/articleview/1884/1/1/

European Commission (2001), White Paper — European transport policy for 2010: time to decide, Office for Official Publications of the European Communities, Luxembourg

International Maritime Organisation (2003), The International Ship and Port Security Code, 2003 Edition

Seising, R.: Die Fuzzifizierung der Systeme: Die Entstehung der Fuzzy Set Theorie und ihrer ersten Anwendungen - Ihre Entwicklung bis in die 70er Jahre des 20. Jahrhunderts. Stuttgart 2005

World Customs Organisation (2007): WCO SAFE Framework of Standards, 2007

Zadeh, L.A.: Information and Control, Fuzzy Sets, 1965, http://www-bisc.cs.berkeley.edu/Zadeh-1965.pdf

Session H2: Scheduling 1

· **Day2: Sep. 16 (Thu.)**

· **Time: 14:40 - 16:00**

· **Chair: Rommert Dekker**

· **Room: Azalea & Lilac, 5F**

A BEAM SEARCH ALGORITHM FOR MINIMIZING RESHUFFLE OPERATIONS AT CONTAINER YARDS

Kun-Chih Wu[1*] and Ching-Jung Ting[2]

[1] Department of Industrial Engineering and Management
Yuan Ze University
135 Yuan-Tung Road
Chung-Li, Taiwan, R.O.C.
e-mail: s968907@mail.yzu.edu.tw

[2] Department of Industrial Engineering and Management
Yuan Ze University
135 Yuan-Tung Road
Chung-Li, Taiwan, R.O.C.
e-mail: ietingcj@saturn.yzu.edu.tw

Abstract: The container relocation at container yards will influence the yard operator efficiency. The problem we tackled is to retrieval all containers from an initial configuration of given number of bays and tiers with minimum number of relocations. There is no new arriving container and only the containers in the same stack above the retrieved container can be relocated. We propose a beam search (BS) algorithm that is similar to the breadth-first branch-and-bound scheme, but only few permitted nodes in each level are kept for the further search. In order to select permitted nodes in each level, a look-ahead heuristic is applied for evaluation. We randomly generate 1920 instances with different configurations to test our BS algorithm. The results show that our BS algorithm can obtain optimal solutions in small size instances and near optimal solutions in larger size instances within effective computation time.

1. INTRODUCTION

With the increasing globalization, the demand of container transportation grows steadily. Terminals have to improve their performance efficiency to compete with others. Many issues are involved in affecting the terminal performance, such as labor utilization, traffic congestion in terminal, transfer time of yard cranes, etc. (Murty *et al*., 2005). Container relocation is a critical issue with respect to the terminal performance, which affects the turn-around time of trucks and of vessels and further impacts the performance of a terminal (Kim and Kim, 1999). The relocation of a container is performed when this container is above another container that is the target to be retrieved immediately. Since relocation is not a value-added but time consuming activity, it should be reduced as many as possible.

Containers are piled up at the container yard in such a manner to increase the space utilization. Block stacking is the most common way for container stowage at container yard (Kim and Hong, 2006), as shown in Figure 1. The numbered block indicates the retrieval order of each container, i.e. the smaller number indicates the higher priority. Though, the block stacking also indicates that only the top container is accessible. On the other hand, for export containers, loading sequence of containers is usually determined before loading to vessel (Lee and Chao, 2009). According to loading sequence, each export container has its own priority to be retrieved. It also implies the relocation occurs when a container with higher priority is underneath another container with lower priority.

Stowage planning is considered to decide the most suitable locations of containers at container yard (Avriel *et al*, 1998). It can be decomposed to two types, the static problem and the dynamic problem, in terms of the arrival/departure manner of containers. In the static problem, the sequence of containers is known in advance, while in the dynamic problem, the containers arrived or departed in a random manner.

For the dynamic problem, stochastic concept is usually adapted to estimate the expected number of relocations. Kim and Kim (1999) dealt with a dynamic relocation problem and presented a formula to interpret the relation between the stack height and the expected number of relocations. Kim *et al*. (2000) considered minimizing expected numbers of relocation at the export container yard. They proposed a dynamic programming model to provide an optimal decision of stack for minimizing expected number of relocations and a decision tree to make a fast decision.

In the real time cases, the exact location for relocating container has to be determined immediately. Therefore, Avriel *et al*. (1998) considered the problem is to minimize the total cost of reshuffles when containers on a vessel have to be loaded and unloaded from/to several ports. The authors proposed a binary programming model as well as a heuristic procedure..

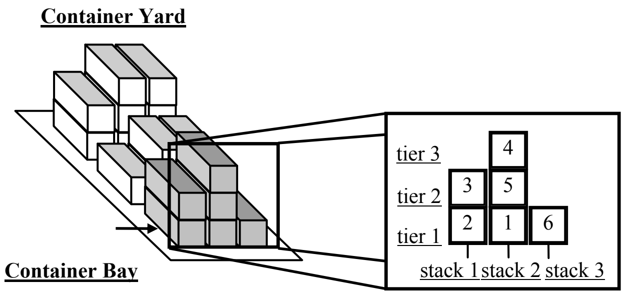


Figure 1. The Illustration of Containers Stacking at the Container Yard

Murty *et al*. (2005) provided a Decision-Support System (DSS) for Hong Kong International Terminals (HIT). The system achieved elastic capacity successfully, almost 50 percent improvement in capacity was obtained. They divided the real world problem into five related subproblems. Reshuffle problem is also an important issue considered in their study, in which a rule called the reshuffle index (RI) was performed to relocate containers. Yang and Kim (2006) addressed a general concept of relocation problem. They proposed both dynamic programming and genetic approach for a static grouped location problem, as well as three rules for the dynamic case.

Kim and Hong (2006) considered a container retrieval problem at the container yard. The objective is to minimize relocations when all containers at a yard-bay are retrieved according to a predetermined precedence. They proposed a branch-and-bound algorithm and a heuristic with respect to reducing the expected number of relocations during retrieval process. The empirical results showed that the error rate between the branch-and-bound and the heuristic is 7.3% in the first case and 4.7% in the later case.Caserta *et al* (2009) continued the study of Kim and Hong (2006). They also handled the container retrieval problem but applied an inspired metaheuristic, called corridor method (CM). The results indicated that CM can reach optimal solution in the small size problem within very short computational time. In this study, we also consider the same problem proposed by Kim and Hong (2006), but we use beam search algorithm to solve this problem.

The remainder of the paper is organized as follows. In section 2, the container reshuffle problem is defined and related heuristics are introduced. The beam search and the branch-and-bound algorithms are introduced in section 3. Section 4 presents the experiment results with large number of instances. Section 5 summarizes the findings in this research.

2. CONTAINER RESHUFFLE PROBLEM

Container relocation problem considered in this paper is to retrieve all containers at a yard-bay. We assume that each container with different retrieval priority and the initial configuration of the yard-bay are known in advance. In addition, we only consider a static problem that means the new arrival container is not allowed. The objective is to retrieve all containers at the yard-bay with minimum number of relocations. The optimal retrieval procedure is performed by determining a stack to relocate a container. Formally, this problem includes following five assumptions:

1) The maximum number of stacks and maximum number of tiers at a container bay are given, and the container only can be relocated within the yard-bay.
2) Every container at the yard-bay has the same size.
3) The initial configuration of the container bay and the retrieved sequence of each container are known in advance.
4) The new arrival container is not allowed during the retrievals.
5) The reshuffle operations only occur while a target container needs to be retrieved. It implies that pre-marshalling is not allowed in this problem.

We use the following notations to describe the problem:

- H : The maximum number of tiers for every stack.
- N : The number of containers in the initial configuration.
- $R(i, j)$: The container at the top of ith stack is relocated to jth stack, where $i, j \in \{1, 2, \ldots, W\}$
- S_i : The set represents the configuration of the ith stack, where $i \in \{1, 2, \ldots, W\}$. The elements in the set indicate the retrieval order of each container, and the last element is on the top of stack.
- W : The maximum number of stacks.
- Y : The set of stacks within the yard bay, where $Y = \{S_1, S_2, \ldots, S_W\}$

An example illustrated in Figure 2 has an initial configuration $Y = \{S_1, S_2, S_3\} = \{\{2, 3\}, \{1, 5, 4\}, \{6\}\}$. The container retrieving is named *target container* henceforth, and the container that blocks the target container retrieving is named *obstructive container*. At the first stage, the target container is container 1. The container 1 cannot be retrieved, because containers 4 and 5 are above it. The relocation is therefore necessary to free the target container. In the first stage, the relocation of container 4 from stack 2 to stack 1 is denoted as $R(2, 1)$. Similar procedure is applied for remainder containers until all containers are retrieved, and the total number of reshuffles in this example is four. The solution can be expressed as all reshuffle operations involved during the retrieval procedure. The solution with four reshuffles can be simply expressed as $\{R(2, 1), R(2, 3), R(1, 3), R(1, 2)\}$.

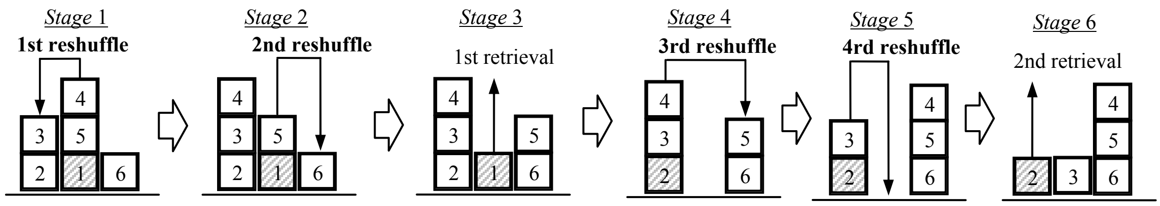


Figure 2: An Example of the Procedure for Retrieving and Reshuffling Containers (H = 3, W = 3)

2.2 Heuristic

There are three heuristics involved in this paper for solving container relocation problem. The first rule, called the lowest position (TLP), was proposed by Zhang (2000). TLP tends to force every stack with a close number of containers, so that average relocation can be reduced. This rule selects a stack that has the fewest containers. Break ties arbitrarily, if two stacks have the same lowest position. For example, in Figure 3 (a), the container 4 has to be relocated to other stacks. The stack 3 has lower position than stack 1, the container 4 then will be relocated to stack 3.

The second rule, called reshuffle index (RI) which computes the number of containers with higher priority than the obstructive container is located, was proposed by Murty *et al.* (2005). Since containers with higher priority underneath the relocated container lead to additional reshuffles. This rule chooses a stack that has lower RI to relocate container. If there are several stacks with the same RI, then break ties arbitrarily. An illustration shown as Figure 3 (b), the container 5 needs to be relocated. The stack 3 has one container with higher priority than container 5, so that stack 3 has lowest RI equal to one. The container 5 should be relocated to stack 3.

The last one is reshuffle index with look-ahead (RIL) that is an extension of reshuffle index rule. The difference between RI and RIL is the rule for breaking ties. In the RIL, the stack for relocating stack is first determined according to RI, and the RIL breaks ties by look-ahead mechanism. The look-ahead rule selects a stack, in which the highest priority of the containers is the lowest. The concept of the look-ahead rule is to avoid additional relocations in near future steps. If a relocated container is above a container that will be retrieved just after the target container, then an extra relocation will occurs in retrieving next container. In Figure 3 (c), the container 4 has to be retrieved, and stack 1 and stack 3 both have equal RI. If container 4 is relocated to stack 1, then one more relocation would be needed when retrieving container 2. According to RIL rule, stack 3 is chosen because the priority of container 3 is lower than the priority of container 2.

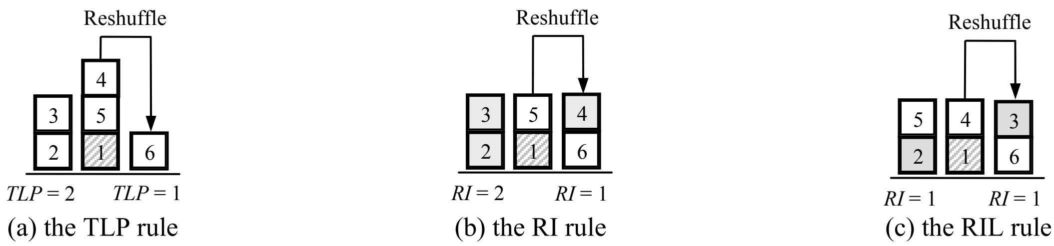


Figure 3. The Heuristic Rules for Relocating a Container

3. BEAM SEARCH

Beam Search (BS) is a heuristic based on breadth-first branch-and-bound algorithm. It was earliest applied in artificial intelligent to deal with the speech recognition (Lowerre, 1976). It has been then employed on the scheduling problem, like the studies conducted by Fox (1983) and Ow and Smith (1988). Ow and Morton (1988) later proposed a novel mechanism for the beam search, called Filter Beam Search, which gives a compromise between computational time and solution quality. They suggested that a rough evaluation resembles a filter is applied first to prune out the nodes, and only α nodes are retained. These α nodes are evaluated again by a caution estimated heuristic and only β nodes are chosen. Sabuncuoglu and Bayiz (1999) applied filter beam search for scheduling problem with makespan and tardiness criteria. A tiny revision for their beam search is that nodes in the first level are evaluated without filter, and a filter is applied in remainder levels.

3.1 Standard Beam Search

BS follows breath-first search strategy in its searching scheme, in which the nodes are searched level by level. Unlike the general tree search scheme that generates every possible node to seek optimal solution, BS only keeps some promised nodes at each level for sprouting the descendants. In the beam search, only β promised nodes at each level are required for the further search. The β is so-called the beam width. The BS reduces a great deal of searching space via limiting the nodes retained at each level. Suppose n is the depth of a tree, the searching space of the BS includes βn solutions against $n!$ solutions contained in the original solution space. However, the optimality can not be confirmed by the beam search, since the optimal solution may be ignored during searching process. The beam search is hence a constructive heuristic.

The promised nodes can be determined simply by estimated rules or heuristics. The nodes with better estimated values have more possibility to gain better results. The concept of beam search is to keep the nodes with higher probability achieving optimal solution during the searching process. The search manner is different from the branch-and-bound that keeps the nodes via a strict restriction, but the beam search realize the concept in an opportunistic manner.

3.2 Beam Search for Container Reshuffle Problem

The decision involved in the container reshuffle problem is to decide which stack to relocate a container. The best sequence of stacks for relocations results in the minimum total number of reshuffles. A node in our beam search represents a partial solution that includes the stacks for relocations from root to current node.

Generally, BS chooses the best β nodes at each level as beam nodes. As illustration in Figure 4, there are four nodes in the level three generated by pervious beam nodes (N_3 and N_4), and two of them generated from the same node (N_3) are chosen as next beam nodes. The search manner concentrates on the partial beams that have better performance. It can increase the possibility to gain better solution via exploiting the partial beams, however it also increases the possibility to fall into the local optimal solution.

In this study, we apply a standard beam search to solve container relocation problem, in which the filter mechanism is not considered. The search begins from an initial configuration, and then the nodes at first level are sprouted out. Every node is evaluated by a given global heuristic to estimate its upper bound. The global heuristic can use one of those three mentioned in section 2.2 to estimate an upper-bound. Only β nodes with lower upper-bound will be selected as beam nodes. The selected beam nodes then sprout out their descendants for the next level. Repeating the procedure until a feasible solution is obtained. An illustration shown as Figure 4, the β of the beam search is set as two. The evaluation function is performed to estimate the upper bound and the search scheme selects the best two descendants to be beam nodes. Only the beam nodes have to generate the descendants at each level, remaining nodes are ignored. The process is terminated at level four, in which a complete solution is obtained. The minimum relocation is four according to the beam nodes at last level.

Let N_0 indicates the root, B is the set of beam nodes, C is the nodes pool of all descendants that generated in the same level. The procedure of beam search is presented as follows:

Beam Search Procedure

Step 1: **[initialization]**

  Set $B = \{N_0\}$

Step 2: **[retrieval procedure]**

  Step 2.1: **For** each node $N \in B$

    Do

      If target container T is on the top of a stack in the bay Y of the node N

        Retrieve the T from Y

    Until there is a container blocked the target container.

  Step 2.2: If a configuration Y in a node $N \in B$ is empty, then go to Step 5.

Step 3: **[branch procedure]**

  Step 3.1: Set the descendants pool $C = \{\emptyset\}$

  For each node $N \in B$

Let S_t is the stack with target container T in node N, and $D = N$

Step 3.2: **For** each stack S in node D

If stack S is not full and $S \neq S_t$

Perform the relocation $R(S_t, S)$

Step 3.3: $C = C \cup D$.

Step 4: **[relocation procedure]**

Step 4.1: For each node $D \in C$

Perform a **Heuristic** to evaluate node D.

Step 4.2: Select best $Min\{\beta, |C|\}$ nodes from C to replace beam nodes in B. Go to Step 2.

Step 5: **[termination]**

Return Min\{number of relocations for node N, $N \in B$\}. Stop the procedure.

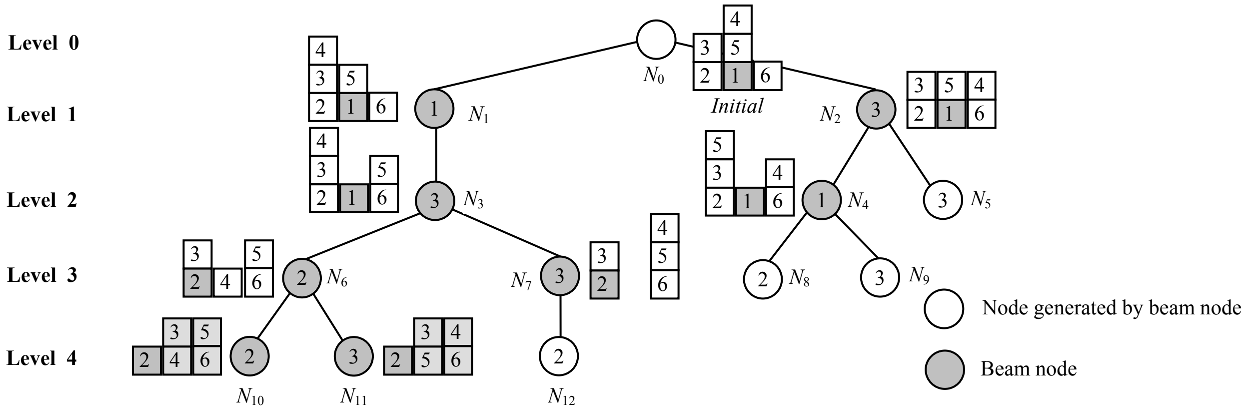


Figure 4. Representation of beam search for container reshuffle problem

3.3 Branch and Bound

We developed a brand-and-bound procedure in this paper as well. The depth-first procedure is applied to reduce memory usage. The node that is sprouted out first will be visited first. Once a feasible solution that has retrieved every container at the yard-bay is found, the procedure will then backtrack to its parent to search another branch. Because the level indicates the number of relocations, if a feasible solution is found at level k, then the further search over level k-1 is not necessary. All generated nodes will be visited to confirm the optimal solution is achieved.

4. EXPERIMENTAL RESULTS

Both the beam search and the branch-and-bound were tested to compare the performance of each other. The three heuristics introduced in section 2.2 were tested as well. All algorithms were coded in Borland C++ Builder 6.0, and run on a PC with AMD Athlon Dual Core +3800 2.0 GHz and 2 GB memory. The testing data are generated randomly for experiments.

4.1 Testing Data

For comparing the performance of algorithms, different sizes of problem were generated. Kim and Hong (2006) mentioned that a yard bay usually contained 4-7 tiers with 6-10 stacks. Hence, without loss of generality, we created problems between 3-8 tiers and 3-10 stacks. We generated 48 problem classes, and each of them has 40 instances, so there are total 1920 instances generated. In each instances, we randomly located containers with different priorities into a yard-bay. In addition, the number of containers (N) is determined by W and H. Given W and H, the total number of slots at the yard-bay is $W \times H$, and the most number of relocations to retrieve one container is $H - 1$. For confirming the feasibility, Eq. (1) ensures that enough space is reserved for relocations of containers.

$$N = W \times H - (H - 1) \tag{1}$$

4.2 Computational Results

We first tested the heuristics with three various rules on each class, and the results are shown as Table 1. The first column and sixth column indicate the number of stack (*W*) with range from 3 to 10. The second column and seventh column indicate the number of tiers (*H*) with range from 3 to 8. The column 3-5 and column 8-10 represent the results of three heuristics. Each cell shows the average number of reshuffles over 40 instances. The results demonstrated that the RIL outperforms both TLP and RI at every class. Because the computational times of the heuristics are quite fast (less than 0.01 seconds), they are not included in the table.

The beam searches with different beam width as well as branch-and-bound algorithm were tested. Ideally, a more accurate estimation leads to a better result of beam search. Therefore, the RIL is embedded in our beam search for global evaluation, due to the better estimation of relocations.The detail results of beam search in each class were shown in Table 3. As the beam width increases, the elapsed time increases linearly, but the improving rate on the average number of reshuffles decreases. The average number of relocations will finally converge to a certain value. An exhaustive search may take too much time but gain little improvement. Hence, the beam search provides a compromise between time and solution quality.

5. CONCLUSION

In this paper, we deal with a container reshuffle problem, in which each container has a particular priority to retrieve. We first proposed a modified heuristic RIL and compared the heuristic with the other two existing heuristics (RI and TLP). Then, we proposed a beam search heuristic to solve the problem, in which the RIL is embedded to provide a global evaluation. We randomly generate 48 problem classes and 40 instances in each class, and there were total 1920 instances generated. The experiments were performed by running those algorithms on generated instances.

The computational results show that the proposed modified heuristic outperforms the other two existing heuristics at every problem class. A branch-and-bound also provided to compare with the beam search. The result shows that the beam search can reach most optimal solution when beam width (β) equal to 15, and all optimal solution can be obtained as beam width equal to 70.

6. REFERENCES

Avriel, M., Penn, M., Shipirer, N. and Witteboon, S. (1998) Stowage Planning for Container Ships to Reduce the Number of Shifts. *Annals of Operations Research, 76:* 55-71.

Caserta, M. and Voß, S. (2009) A Corridor Method-based Algorithm for the Pre-marshalling Problem. *Lecture Notes in Computer Science, 5484:* 788-797.

Caserta, M., Schwarze, S. and Voß, S. (2009) A New Binary Description of the Blocks Relocation Problem and Benefits in a Look Ahead Heuristic. *Lecture Notes in Computer Science, 5482:* 37-48.

Fox, M. S. (1983) *Constraint-Directed Search: A Case Study of Job-Shop Scheduling*, PhD Dissertation, Computer Science Department, Carnegie-Mellon University.

Kim, K. H. and Hong, G. P. (2006) A Heuristic Rule for Relocating Blocks. *Computers & Operations Research, 33:* 940-954.

Kim, K. H. and Kim, H. B. (1999) Segregating Space Allocation Models for Container Inventories in Port Container Terminals. *International Journal of Production Economics, 59:* 415-423.

Kim, K.H. and Kim, H.B. (2002) The Optimal Sizing of the Storage Space and Handling Facilities for Import Containers. *Transportation Research Part B, 36:* 821-835

Kim, K. H., Park, Y. M. and Ryu, K. R. (2000) Deriving Decision Rules Locate Export Containers in Container Yards. *European Journal of Operational Research, 124:* 89-101

Lowerre, B.T. (1976) The HARPY Speech Recognition System. Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, PA.

Lee, Y. and Chao, S-L. (2009) A Neighborhood Search Heuristic for Pre-marshalling Export Containers. *European Journal of Operational Research, 196:* 468-475.

Murty, K.G., Liu J., Tseng, M. M., Leung, E., Lai, K-K. and Chiu, W.C. (2005) Hong Kong International Terminals Gains Elastic Capacity Using a Data-intensive Decision Support System. *Interfaces, 35:* 61-75.

Ow, P. S. and Morton, T. E. (1988) Filtered Beam Search in Scheduling. *International Journal of Production Research 26:* 35-62

Ow, P. S. and Morton, T. E. (1989) The Single Machine Early/Tardy Problem. *Management Science, 35:* 177-191.

Ow, P. S. and Smith, S. F. (1988) Viewing Scheduling as an Opportunistic Problem-solving Process. *Annals of Operations Research, 12:* 85-108.

Sabuncuoglu, I. and Bayiz, M. (1999) Job Shop Scheduling with Beam Search. *European Journal of Operational Research, 118:* 390-412.

Sabuncuoglu, I., Gocgun, Y. and Erel, E. (2008) Backtracking and exchange of information - Methods to enhance a beam search algorithm for assembly line scheduling. *European Journal of Operational Research, 186:* 915-930.

Valente, J. M. S. and Alves, R. A. F. S. (2005) Filtered and Recovering Beam Search Algorithms for the Early/Tardy Scheduling Problem with No Idle Time. *Computers & Industrial Engineering, 48:* 363–375.

Yang, J.H. and Kim, K. H. (2006) A Grouped Storage Method for Minimizing Relocations in Block Stacking Systems. *Journal of Intelligent Manufacturing, 17:* 453-463.

Zhang, C. (2000) *Resource Planning in Container Storage Yard*, PhD Dissertation, Department of Industrial Engineering, The Hong Kong University of Science and Technology.

Table 1. Average Numbers of Reshuffles of Evaluation Heuristics in Each Class

class		TLP	RI	RIL	class		TLP	RI	RIL
W	*H*				*W*	*H*			
3	3	3.58	3.42	3.4	7	3	11.50	10.10	9.88
3	4	6.67	6.10	5.95	7	4	19.40	20.10	18.80
3	5	10.60	9.80	9.45	7	5	33.00	30.90	28.30
3	6	15.40	13.60	13.20	7	6	45.80	45.00	41.30
3	7	20.20	18.10	17.30	7	7	66.10	59.80	54.50
3	8	27.20	24.10	22.90	7	8	82.60	76.60	68.70
4	3	5.67	5.03	4.92	8	3	11.60	11.90	11.70
4	4	10.50	9.05	8.80	8	4	22.90	20.90	20.00
4	5	16.30	14.50	13.70	8	5	37.60	34.70	32.10
4	6	23.20	19.10	17.90	8	6	55.00	50.60	46.50
4	7	33.40	28.90	27.60	8	7	79.20	68.40	61.60
4	8	44.80	37.90	34.60	8	8	93.90	88.90	79.50
5	3	6.95	5.90	5.75	9	3	14.10	12.40	12.10
5	4	14.40	12.20	11.80	9	4	24.80	24.70	23.60
5	5	21.00	18.10	17.40	9	5	44.50	37.50	35.10
5	6	31.80	25.60	24.10	9	6	61.50	55.70	50.20
5	7	46.00	36.30	33.70	9	7	88.50	76.70	68.80
5	8	61.60	49.80	44.50	9	8	109.00	97.90	87.40
6	3	8.95	7.92	7.80	10	3	14.90	14.60	14.40
6	4	16.00	13.20	12.80	10	4	28.90	26.20	24.70
6	5	26.90	22.60	21.40	10	5	47.90	41.20	38.50
6	6	41.20	32.60	30.10	10	6	67.00	60.20	54.40
6	7	56.90	45.50	41.70	10	7	101.00	85.10	75.80
6	8	75.00	57.20	53.90	10	8	121.00	111.00	96.70
					avg		39.708	34.950	32.069

Table 2. Average Numbers of Reshuffles of Beam Search with Varied Beam Width and the Branch-and-Bound

class		Beam Search												B&B	
		BW=5		BW =10		BW =15		BW =20		BW =50		BW =70		Opt.	
W	*H*	Res	Time	Res	Time	Res	Time	Res	Time	Res	Time	Res	Time	Res	Time
3	3	**3.38**	0.000	**3.38**	0.000	**3.38**	0.000	**3.38**	0.000	**3.38**	0.000	**3.38**	0.000	**3.38**	0.000
3	4	**5.67**	0.000	**5.67**	0.000	**5.67**	0.000	**5.67**	0.000	**5.67**	0.000	**5.67**	0.000	**5.67**	0.000
3	5	**8.40**	0.000	**8.40**	0.000	**8.40**	0.000	**8.40**	0.001	**8.40**	0.001	**8.40**	0.001	**8.40**	0.001
3	6	**11.50**	0.001	**11.50**	0.001	**11.50**	0.001	**11.50**	0.002	**11.50**	0.003	**11.50**	0.004	**11.50**	0.009
3	7	15.10	0.001	15.10	0.001	**15.00**	0.002	15.10	0.003	**15.00**	0.006	**15.00**	0.008	**15.00**	0.060
3	8	19.00	0.002	18.90	0.003	18.80	0.004	18.80	0.005	18.70	0.010	**18.60**	0.014	**18.60**	1.510
4	3	**4.85**	0.000	**4.85**	0.000	**4.85**	0.000	**4.85**	0.000	**4.85**	0.000	**4.85**	0.000	**4.85**	0.000
4	4	**8.43**	0.000	**8.43**	0.001	**8.43**	0.001	**8.43**	0.001	**8.43**	0.002	**8.43**	0.003	**8.43**	0.007
4	5	12.30	0.001	**12.20**	0.002	**12.20**	0.002	**12.20**	0.003	**12.20**	0.006	**12.20**	0.007	**12.20**	0.241
4	6	15.80	0.002	15.70	0.003	15.70	0.004	15.70	0.005	**15.60**	0.012	**15.60**	0.016	**15.60**	11.000
4	7	23.00	0.003	22.90	0.006	22.80	0.008	22.70	0.010	22.60	0.023	22.60	0.031	–	–
4	8	28.90	0.005	28.40	0.009	28.20	0.013	28.10	0.017	28.00	0.038	27.90	0.052	–	–
5	3	**5.75**	0.000	**5.75**	0.000	**5.75**	0.001	**5.75**	0.000	**5.75**	0.002	**5.75**	0.002	**5.75**	0.002
5	4	**11.00**	0.001	**11.00**	0.002	**11.00**	0.002	**11.00**	0.002	**11.00**	0.006	**11.00**	0.008	**11.00**	0.308
5	5	15.70	0.002	15.70	0.004	**15.60**	0.005	**15.60**	0.007	**15.60**	0.014	**15.60**	0.019	**15.60**	40.900
5	6	21.30	0.004	21.30	0.007	21.20	0.009	21.20	0.013	21.10	0.030	21.10	0.040	–	–
5	7	28.30	0.006	28.10	0.012	28.10	0.017	27.90	0.022	27.80	0.052	27.80	0.070	–	–
5	8	37.50	0.011	37.00	0.020	36.90	0.029	36.90	0.038	36.50	0.089	36.40	0.121	–	–
6	3	**7.65**	0.000	**7.65**	0.002	**7.65**	0.002	**7.65**	0.002	**7.65**	0.004	**7.65**	0.005	**7.65**	0.034
6	4	12.10	0.002	**12.00**	0.003	**12.00**	0.004	**12.00**	0.005	**12.00**	0.012	**12.00**	0.016	**12.00**	11.200
6	5	19.40	0.004	19.40	0.007	19.40	0.011	19.40	0.013	19.40	0.030	19.30	0.041	–	–
6	6	26.50	0.007	26.30	0.013	26.30	0.019	26.30	0.025	26.10	0.057	26.10	0.079	–	–
6	7	35.80	0.012	35.60	0.023	35.30	0.034	35.20	0.045	35.00	0.107	34.90	0.148	–	–
6	8	44.30	0.018	44.20	0.036	43.80	0.052	43.50	0.068	43.20	0.162	43.20	0.226	–	–
7	3	**8.95**	0.001	**8.95**	0.002	**8.95**	0.003	**8.95**	0.004	**8.95**	0.007	**8.95**	0.011	**8.95**	0.834
7	4	15.50	0.003	15.50	0.005	15.50	0.008	15.50	0.010	15.50	0.023	15.50	0.031	–	–
7	5	21.60	0.005	21.50	0.011	21.40	0.016	21.40	0.020	21.40	0.049	21.40	0.066	–	–
7	6	31.60	0.011	31.30	0.022	31.30	0.032	31.20	0.042	31.00	0.099	31.00	0.136	–	–
7	7	40.50	0.019	40.20	0.036	40.00	0.053	39.80	0.069	39.50	0.165	39.40	0.229	–	–
7	8	51.70	0.031	51.00	0.060	50.80	0.089	50.50	0.116	50.00	0.277	49.90	0.383	–	–
8	3	**9.72**	0.002	**9.72**	0.003	**9.72**	0.004	9.72	0.005	**9.72**	0.011	**9.72**	0.015	**9.72**	6.510
8	4	17.90	0.004	17.90	0.008	17.90	0.012	17.90	0.016	18.00	0.035	18.00	0.048	–	–
8	5	25.60	0.009	25.50	0.017	25.50	0.025	25.40	0.033	25.40	0.079	25.40	0.107	–	–
8	6	36.80	0.017	36.40	0.033	36.20	0.048	36.20	0.064	36.00	0.153	36.00	0.209	–	–
8	7	46.20	0.028	45.60	0.054	45.60	0.080	45.60	0.105	45.20	0.251	45.10	0.345	–	–
8	8	59.30	0.048	58.50	0.091	58.50	0.136	58.00	0.179	57.50	0.427	57.30	0.593	–	–
9	3	**11.40**	0.002	**11.40**	0.004	**11.40**	0.007	**11.40**	0.009	**11.40**	0.018	**11.40**	0.025	**11.40**	147.000
9	4	19.20	0.006	19.20	0.013	19.10	0.018	19.10	0.024	19.10	0.055	19.10	0.075	–	–
9	5	28.90	0.013	28.90	0.025	28.80	0.038	28.80	0.050	28.70	0.118	28.70	0.161	–	–
9	6	40.80	0.025	40.60	0.048	40.50	0.072	40.30	0.095	40.00	0.227	40.00	0.316	–	–
9	7	52.80	0.043	52.40	0.084	52.20	0.124	52.00	0.162	51.70	0.396	51.50	0.547	–	–
9	8	68.10	0.073	67.20	0.139	67.20	0.209	66.90	0.277	66.40	0.665	66.00	0.910	–	–
10	3	11.90	0.003	11.90	0.006	11.90	0.009	11.90	0.011	11.90	0.024	11.90	0.033	–	–
10	4	22.50	0.009	22.40	0.017	22.40	0.025	22.40	0.034	22.40	0.079	22.40	0.108	–	–
10	5	32.10	0.019	31.90	0.036	31.90	0.054	31.90	0.072	31.80	0.170	31.80	0.235	–	–
10	6	44.70	0.035	44.40	0.069	44.20	0.101	44.10	0.134	43.90	0.321	44.00	0.451	–	–
10	7	59.20	0.065	58.60	0.127	58.40	0.184	58.00	0.243	57.70	0.588	57.50	0.817	–	–
10	8	73.00	0.097	72.20	0.189	72.10	0.279	71.70	0.367	71.40	0.896	71.20	1.240	–	–
avg		30.349	0.023	30.124	0.045	30.049	0.066	29.959	0.087	29.805	0.210	29.752	0.292		

A GA-BASED APPROACH FOR CONTAINER UNLOADING SCEDULING PROBLEM WITH MH'S STABILITY CONSTRAINT

Kyuhyeon Shin[1], and Taesik Lee[1*]

[1]Department of Industrial and System Engineering
Korea Advanced Institute of Science and Technology
Guseong-dong, Yuseong-gu
Daejeon, Korea
e-mail: taesik.lee@kaist.edu

Abstract: Mobile Harbor is a movable, floating structure with a container crane on board. Containers are picked up from a containership and placed on the deck of a Mobile Harbor, and vice versa for the loading process. In planning container transfer tasks for the container crane, two factors must be considered: task completion time and the stability of the Mobile Harbor. During the un/loading process, weight distribution of container stack on the deck of the Mobile Harbor affects the stability of the floating structure. A target container and its destination on the deck of the Mobile Harbor must be determined to satisfy stability conditions. Also required is a reduction of the crane's movement on the Mobile Harbor to minimize the task completion time. These two factors are related to each other and thus need to be solved together. In this paper, we present a Genetic Algorithm approach to solving the problem.

1. INTRODUCTION

The Mobile Harbor (MH) is a new port system concept developed at KAIST. The MH is a movable, floating structure with a container crane on board; it can unload containers from a container ship anchored out in the sea and place them on the deck of a MH, and vice versa for the loading process. A container ship does not have to call at a land-berth for loading and unloading if a MH can serve the ship.

The MH loading and unloading operation is different from that of an existing land-based port because of the operational characteristics unique to a MH. The operation scheduling for a MH system must cope with these characteristics. Especially, scheduling a loading and unloading sequence of the MH crane is very different from the scheduling of a quay crane. For example, minimizing the number of lateral motions – i.e. moves from stem to stern – of the on-board crane is an important factor to improving the loading/unloading productivity. Stability of a MH during an unloading and loading process is another factor to consider, which does not exist in case of the quay crane scheduling.

While the container unloading scheduling is an important problem for a MH operation, it is not crucial to a quay crane operation of a land-based port. As a result, there is little previous study that is directly relevant to the problem we are undertaking. On the other hand, much research has been done in the area of quay crane scheduling and container yard operations. For example, Kim and Park (2004) and Monica *et al.* (2006) studied a Quay Crane Scheduling Problem (QCSP) using a mixed integer programming model. Sammarra *et al*. (2007) suggested a tabu search heuristics for QCSP. Heuristic methods for routing yard-side equipment are presented in Kim and Kim (2003). Although these studies provide a useful modeling framework for the MH unloading problem, to reflect unique characteristics of a MH requires significant modification of such models. Thus, the focus of this paper is to propose a formulation and a solution method for the container unloading scheduling by incorporating unique characteristics of a MH.

In this paper, the on-board crane movement schedule for unloading containers from a container ship to a MH is generated. In particular, when each container is placed on the deck of the MH, the center of gravity (CG) of the MH with the containers stacked on its deck is restricted to the proximity of the original CG of the MH, in order to ensure its stability. This schedule should also minimize the lateral crane movements, to minimize the loss of unloading productivity. To achieve these objectives, we define this problem with a mathematical model and solve it with a Genetic Algorithm (GA).

2. PROBLEM & MATHEMATICAL MODELING

2.1 Problem Description

A MH unit can unload up to 240 TEU (Twenty-foot Equivalent Unit) containers from a container ship to the deck of the MH. In other words, the maximum carrying capacity of a MH unit is 240 TEU. In this paper, all the containers are assumed to be forty-foot containers (FEU, Forty-foot Equivalent Unit). 240 TEU containers are the equivalent of 120 FEUs, and they are stacked on the deck of a MH unit in ten rows, four bays and three tiers; that is, 40 FEUs are stored in each tier (Fig. 1). In addition, the work range of the MH crane is limited due to the length of a MH unit. Thus, the MH crane can unload containers on a container ship that are stacked in a series of four bays with respect to the MH

Figure 1. A MH and its deck

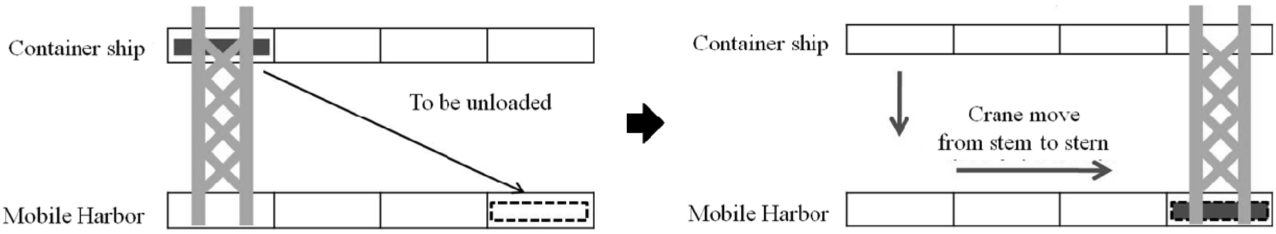


Figure 2. The on-board crane movement from stem to stern

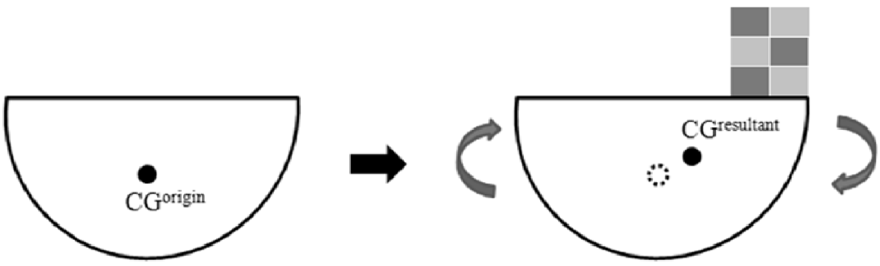


Figure 3. A change in CG of a ship caused by containers

unit's docking position.

Since there is no separate equipment on a MH unit for its deck operation – i.e. handling containers on the deck –, the MH crane must transfer containers from a container ship to a desired location on the deck of the MH directly. This additional move by the crane increases the workload of the on-board crane, compared to a quay crane, and these moves need to be minimized to increase the productivity of the MH. Especially, the crane movement from stem to stern, or the reverse, makes the MH list and negatively affects the stability, and thereby decreases the productivity of the MH. Therefore, it is essential to minimize crane movements during the unloading operation (Fig. 2). In addition, the weight distribution of containers stacked on the deck of a MH unit affects the stability of the MH. If containers are stacked on only one side of a MH unit, a MH will list. In the worst case, the container stack may even collapse and get damaged. So, CG of the MH with the stacked containers must reside within a safe limit (Fig. 3).

The followings are assumptions in this problem.

- 40% of containers to be unloaded from a container ship are empty, and their weight is four tons; the weight of the non-empty containers is uniformly distributed in a range from 10 to 30 tons.
- The number of possible locations, m, at which containers are placed on the deck of t MH unit is a multiple of 40. 40 is the number of container locations in a single tier. This implies that a lower tier is always filled up. For example, if the number of containers to be unloaded is 60, the number of possible unloading locations m is 80, 40 of which are from the first tier and the other 40 from the second tier. Out of 40 locations in the second tier, 20 are used. No container is stacked in the third tier.
- When docked for an unloading operation, a MH unit is always perfectly aligned with a container ship. That is, transferring containers from a bay of a container ship to a corresponding bay of a MH does not require any lateral crane movement along the direction from stem to stern.

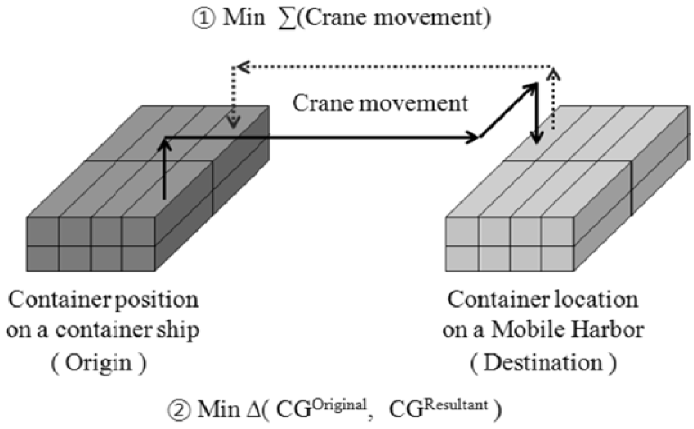


Figure 4. Description of the objective of this problem

With these assumptions, this problem is to find an unloading sequence of n containers and their destination locations on the deck of a MH unit that minimizes a weighted sum of unloading time and the deviation of the resultant CG from the original CG of an empty MH (Δ_{CG}) (see Figure 4).

2.2 Mathematical Model

To construct a mixed integer programming model for this problem, we follow a basic formulation proposed by Kim an Park (2004) in their QCSP. The followings are the notations used in this paper:

$\mathrm{C} = \{1, \cdots, n\}$	The set of containers to be unloaded. n is the number of containers.
$\mathrm{L} = \{1, \cdots, m\}$	The set of candidate locations on the deck of a MH unit at which a container can be placed. m is the total number of candidate locations.
0, T	0 is a dummy location on a MH to indicate a starting position of the on-board crane. T is a dummy container to define the end of the unloading operation.
t_{ij}	The time for the on-board crane to move a container i on a container ship to its destination location j on the MH. Likewise, it is the time for the crane to move from a location i on the MH to a container j on a container ship for a pick up.
LU_j	The set of locations above location 'j' on a MH.
CU_i	The set of containers above container 'i' on a container ship.
k	The number of hatch covers of a container ship within the four-bay work range.
$\mathrm{H_k} = \{1, \ldots, h_k\}$	The set of containers stored in the hold area under a hatch cover k of a container ship.
$\mathrm{G_k} = \{1, \ldots, g_k\}$	The set of containers stored in the deck area above a hatch cover k of a container ship.

The followings are decision variables.

x_{ij}	If the container 'i' on a container ship is transferred to the location 'j' of a MH, or if the crane moves from location 'i' on the MH to the position of the container 'j' on a ship, 1. 0 otherwise. The set of non-zero x_{ij} completely describes each container's destination and their unloading sequence.
Y	Total unloading operation time.
D_i	The time when the crane reaches the position of the container 'i' on a container ship or the location 'i' on a MH.
Δ_{CG}	Distance between CG of the empty MH and new CG when the unloading operation is completed.

A mixed integer programming model is shown below.

$$\text{Minimize} \quad \alpha \mathrm{Y} + \beta \Delta_{\mathrm{CG}} \tag{1}$$

$$\sum_{i \in \mathrm{C}} x_{0i} = 1 \tag{2}$$

$$\sum_{j \in \mathrm{L}} x_{j\mathrm{T}} = 1 \tag{3}$$

$$\sum_{i \in \mathrm{C}} x_{ij} \leq 1, \qquad j \in \mathrm{L} \tag{4}$$

$$\sum_{j \in \mathrm{L}} x_{ij} = 1, \qquad i \in \mathrm{C} \tag{5}$$

$$\sum_{i \in \mathrm{C} \cup \{\mathrm{T}\}} x_{ji} = \sum_{i \in \mathrm{C}} x_{ij} , \qquad j \in \mathrm{L} \tag{6}$$

$$\sum_{j \in \mathrm{L} \cup \{0\}} x_{ji} = \sum_{j \in \mathrm{L}} x_{ij} , \qquad i \in \mathrm{C} \tag{7}$$

$$\mathrm{Y} \geq \mathrm{D}_i , \qquad \forall i \in \mathrm{C} \tag{8}$$

$$\mathrm{D}_j \leq \mathrm{D}_p , \qquad j \in \mathrm{L} \text{ and } p \in \mathrm{LU_j} \tag{9}$$

$$\mathrm{D}_i \leq \mathrm{D}_p , \qquad i \in \mathrm{C} \text{ and } p \in \mathrm{CU_j} \tag{10}$$

$$\mathrm{D}_i \leq \mathrm{D}_p , \qquad i \in \mathrm{G}_k \text{ and } p \in \mathrm{H}_k \tag{11}$$

$$\mathrm{D}_j - \mathrm{D}_i + M x_{ji} \leq M - t_{ji} , \qquad i \in \mathrm{C} \text{ and } j \in \mathrm{L} \cup \{0\} \tag{12}$$

$$\mathrm{D}_i - \mathrm{D}_j + M x_{ij} \leq M - t_{ij} , \qquad j \in \mathrm{L} \text{ and } i \in \mathrm{C} \cup \{\mathrm{T}\} \tag{13}$$

$$x_{ip} = 0 , \qquad i \in \mathrm{C}, p \in \mathrm{C} \cup \{\mathrm{T}\} \tag{14}$$

$$x_{pj} = 0 , \qquad j \in \mathrm{L}, p \in \mathrm{L} \cup \{0\} \tag{15}$$

$$x_{0\mathrm{T}} = 0, \qquad x_{\mathrm{T}0} = 0 \tag{16}$$

$$x_{ii} = 0 , \qquad i \in \mathrm{C} \cup \mathrm{L} \tag{17}$$

The objective function (1) is a weighted sum of a total unloading time and a distance between CG of an empty MH and of the loaded MH. Note that we ignore the interim CG during its operation and only evaluate the resultant CG after the unloading operation is completed. In the heuristic model presented in the next section, however, we evaluate CG every time a container is loaded on a MH to maintain CG to be within a desired limit throughout the unloading operation. The weight α and β in the objective function are both set to 1.

Constraints (2) and (3) define a starting task and a final task of the unloading operation, respectively. Constraints (4) ~ (7) are routing constraints. Constraint (4) is expressed as inequality because there may be some locations on a MH where no container is placed when a tier is not fully populated with containers. Constraint (8) defines a total unloading time for all containers. Constraints (9) ~ (11) are needed to avoid infeasible solutions that have invalid unloading sequences. Constraint (9) prevents a container from being placed at a location in a higher tier if the corresponding location in its lower tier is empty. Constraint (10) states that a container stacked below another container on a container ship cannot be processed before the one above is processed. Constraint (11) requires that, to unload containers stored in the hold area under a hatch cover, all the containers stored above the hatch cover must be unloaded first so that the

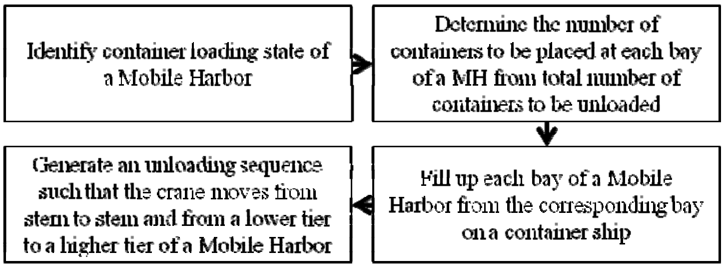


Figure 5. A diagram of a rule-based heuristic method

hatch cover can be cleared. Constraints (12) and (13) set relations between two successive movements in terms of their move completion time. *M* is a large number. Last of all, constraints (14) ~ (17) are ancillary to model.

3. METHOD

While the mathematical model assures an optimal solution, its computational complexity makes it difficult to solve beyond a small sized problem. Even a problem instance with only a few containers takes very long to obtain a solution. As an alternative, a heuristic approach is needed, for problems with a large number of containers, to obtain a near optimal solution in a short period of time. In this section, we present a GA to generate unloading sequences that improve the productivity and the stability of a MH. For comparison, also presented is a local search method, a Hill-Climbing method, which has been suggested in a recent study by Shin and Lee (2010).

3.1 Rule-based heuristic

A rule-based heuristic method is suggested in Lee *et al*. (2010). A rule-based heuristic method generates an unloading sequence that minimizes the crane movement in the stem-to-stern direction. It is based on a simple rule: a container in a bay b^c of a container ship must be moved to an unloading location in the corresponding bay, say bay b^{MH}, of a MH. If no location is available in the b^{MH}, then the container is moved to a location in a bay closest to bay b^c. Figure 5 shows a process of generating an unloading sequence with the above mentioned rule. This method makes the problem simple as the unloading sequence determines unloading locations. That is, if an unloading sequence is given, unloading locations for all containers are automatically determined. Thus, the solution space explored by this heuristic method is more limited than the solution space for other methods presented in this paper.

3.2 Hill-Climbing

A Hill-Climbing is one of local search methods. So the quality of a solution obtained from a Hill-Climbing method depends much on the quality of an initial solution. Shin and Lee (2010) used a Hill-Climbing method for the MH crane scheduling problem, and they used an initial solution generated by the above mentioned rule-based heuristic method. This choice of an initial solution is justified by the fact that the rule-based heuristic method provides the best solution for problems where an unloading sequence is tied to unloading locations. As the Hill-Climbing method is expanding its search in a solution space where an unloading sequence and unloading locations are not tied together, this initial solution seems to offer a reasonable starting point for the Hill-Climbing method.

For the Hill-Climbing search, an initial solution is encoded by a chromosome, as in a GA. The chromosome consists of two sets of orders (see Fig. 6). To the chromosomes, three mutation operators are applied to conduct a Hill-Climbing search. All mutants at each generation are evaluated, and the best change among all mutants is passed on to the next generation. More detailed discussion on the structure of this chromosome and a mutation process applied to the chromosomes are presented in section 3.3 since the same chromosome structure and mutation operation is used in the GA developed in this study.

In general, a Hill-Climbing has a disadvantage that it can only get a local optimum close to an initial solution. For this reason, a search method that can explore a broad solution space is needed for an improved optimality of a solution. In the next section, we present a GA that aims to minimize operation time and improve the stability of a MH during its unloading operation.

3.3 Genetic Algorithm (GA)

In the GA developed in this study, the same chromosome structure as the one described in the Hill-Climbing method is used. A chromosome consists of two sets of orders; one represents an unloading sequence, and the other represents locations on a MH unit to which containers will be moved and placed. (Fig. 6) So, a chromosome has 2*n* genes, where *n*

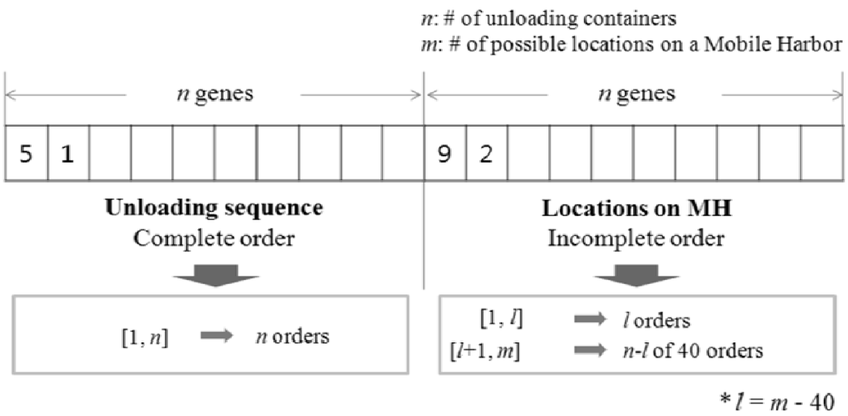


Figure 6. The structure of a chromosome

is the number of containers to be unloaded. The first *n* genes are a set of unloading sequences, and the latter *n* genes are locations of containers on the deck of a MH.

The first part of the chromosome, size of *n* genes, is populated with numbers between 1 and *n* as it indicates the unloading sequence. Thus, we call this a "complete order". On the other hand, *n* genes of the latter part are selected from numbers between 1 and *m*, where *m* is the number of possible locations on the deck of a MH. Note that *m* is always greater than or equal to *n* and is a multiple of 40. It should also be noted that one of the assumptions in the problem states that a lower tier on the deck of a MH is completely filled if there are enough number of containers to fill the lower tier. For example, when there are 60 containers to unload, the first tier of a MH will be filled with 40 containers, and then the remaining 20 containers will be placed somewhere in the second tier. For the chromosome, this means that *n* genes (n = 60, in the example here) of the latter part of the chromosomes will include 1 to 40 and another twenty numbers between 41 and 80. Due to this feature, we call this part of the chromosome an "incomplete order" to indicate that a part of the genes contain an incomplete order.

An evaluation function adopted is divided into two parts: the operation time, and the stability of a MH when the unloading operation is completed. A simple sum of the two components is used as the evaluation function. It is equivalent to setting α and β to 1. The operation time is calculated by the sum of the times of all the crane moves. The stability is measured by the deviation of CG of a MH as containers are placed and stored on the deck of it. The stability is assessed throughout the entire unloading sequence. If the deviation of CG at any point during an unloading sequence gets greater than a desired threshold value, a penalty is given to the chromosome.

Initial populations are generated at random, with a constraint that the above-stated structure of each chromosome must be maintained. Selection uses a tournament selection. A tournament selection randomly picks two chromosomes and selects the better chromosome until the intermediate population is filled.

A crossover operation is modified due to the unique structure of the proposed chromosome. Since a chromosome is divided in two parts, a typical crossover operation is not allowed in that the first and second part of the chromosome should not be mixed. Thus, a crossover operation should be carried out within each part of the chromosome. Also considered is that a change as a result of a crossover operation in one part of the chromosome – the unloading sequence part or the location part – is reflected in the other part of the chromosome accordingly. That is, crossover points selected in one part of a chromosome are also selected on the corresponding location of the other part (Fig. 7).

Special mutation operators are also defined to deal with the unusual chromosome structure. The three mutation operators are a "sequence-exclusive operator", a "location-exclusive operator", and a "sequence-location operator". A sequence-exclusive operator mutates a randomly selected gene in only the first part of a chromosome. A location-exclusive operator generates a mutant in only the second part. A sequence-location operator creates a mutant in the first part of a chromosome and another mutant on a corresponding gene in the latter part of the chromosome. The composition of mutation operators is for information exchange and preservation. In case of the first part of a chromosome, that has the complete order, a value of a created mutant always corresponds to a gene within the order. Thus, a mutation results in an exchange of their positions. For mutations in the second part, where genes have the incomplete order, this position exchange occurs only if a mutant takes a value that is already taken by another gene in the order. When a mutant in the location part takes a value that does not exist in the second part of a current chromosome, different types of handling are required depending on whether the mutant is part of the complete-order potion of the incomplete order or not. If it is, to maintain the complete-order structure, one gene is selected from the incomplete-order portion and its value is replaced by the original value of the mutant gene. This is illustrated in Figure 8. If the mutant is part of the incomplete-portion, this mutation is simply allowed.

Due to a large solution space of the problem, we adopt two strategies. Firstly, elitism is accepted to expedite the search process. The best two solutions of the current generation are directly transferred to the next generation by

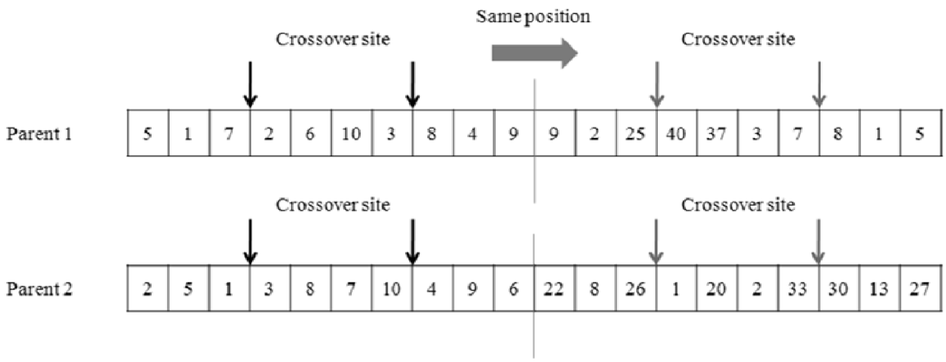


Figure 7. Selection of crossover sites within one part of a chromosome and its duplication in the other part

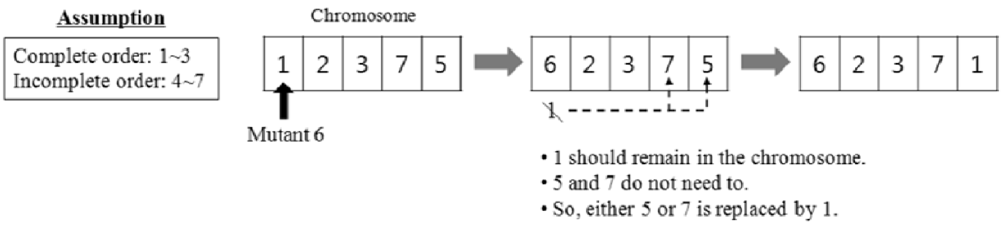


Figure 8. Description of exception handling in the mutation operation

replacing two worst performing solutions of the next generation. Second, a repair function is introduced to eliminate infeasible solutions. After a crossover and mutation operation, a resulting chromosome may become an infeasible solution. A chromosome is infeasible when (1) the container i stacked below the container j on a container ship is to be unloaded earlier than container j, (2) the container i stacked beneath a hatch cover, on which the container j is stacked, is to be unloaded earlier than the container j, and (3) a container is to be placed at the location k on the MH before the location l, which is underneath the location k has a container. Infeasible solutions tend to slow the search process. A repair function helps to cope with this problem. The repair function changes the unloading sequence of an infeasible chromosome to make it a feasible solution. While it is expected to speed up the search process and improve the quality of a solution, it has disadvantages too. It may cause premature convergence by artificially biasing the search process. As a consequence, it may not allow a search for global optimal solutions. In spite of these drawbacks, we adopted the repair function, considering a very large solution space of the problem.

A number of experiments are conducted by varying population size, generation number and the probability of crossover and mutation. Levels of parameters used in the experiments are shown in Table 1.

Table 1. Levels of parameters of the GA used in experiments

Population size	Generation number	Probability of crossover	Probability of mutation
500	1000	0.3	0.001
1000	1500	0.5	0.005
1500	2000	0.6	

4. COMPUTATIONAL RESULTS

Performance of the proposed GA method is analyzed by comparing its outcome with other search method. It is also of our interest to determine whether the repair function defined in this study is effective. Thus, we first compare unloading operation time and the stability values obtained from the GA method with the repair function to the values from the GA method without the repair function. Secondly, we compare the performance of the solutions by the GA method to the Hill-Climbing method.

4.1 Effectiveness of the Repair Function

Crossover and mutation operations in the proposed GA method usually produce an infeasible solution. The repair function introduced in the GA method transforms an infeasible solution to a feasible solution by manipulating part of the unloading sequence. When a repair function is not used, we still penalize an infeasible solution. A penalty of 100 is given to each violation instance of the three infeasibility conditions. We compare the quality of solutions obtained by the GA with the repair function to those obtained by the GA without the repair function. Table 2 summarizes the results from nine experiments.

Table 2. Results from the comparison between the GA with and without the repair function

Experiment sets		GA with the repair function			GA without the repair function		
Number	**# of containers**	**Evaluation value**	**Operation time (min)**	Δ_{CG}**(m)**	**Evaluation value**	**Operation time (min)**	Δ_{CG}**(m)**
1	**50**	**188.73 18**	**8.03**	**0.7**	**191.59 18**	**8.34**	**3.3**
2	**60**	**204.84 20**	**4.34**	**0.5**	**206.42 20**	**4.34**	**2.1**
3	**95**	**314.12 31**	**2.22**	**1.9**	**522.75 51**	**5.13**	**7.6**
4	**100**	**295.29 29**	**0.59**	**4.7**	**743.00 73**	**1.87**	**11.1**
5	**115**	**436.08 43**	**2.38**	**3.7**	**1065.39 1**	**052.60**	**12.8**
6	**120**	**355.93 35**	**2.53**	**3.4**	**683.16 67**	**0.14**	**13.0**
7	**120**	**460.10 45**	**8.20**	**1.9**	**967.4 95**	**8.32**	**9.1**
8	**120**	**510.77 50**	**9.17**	**1.6**	**3421.67 3**	**412.14**	**9.5**
9	**120**	**526.64 52**	**4.54**	**2.1**	**1533.31 1**	**524.66**	**8.7**

As seen from Table 2, the repair function significantly improves the quality of solutions. In fact, for the experiments 3-9, the GA without the repair function cannot even find a feasible solution within the given number of generations.

The best results of GA with the repair function are achieved when the population size is 500, the number of generation is 1000, and the probabilities of crossover and mutation are 0.3 and 0.005, respectively.

4.2 Comparison with the Hill-Climbing

Solutions obtained from the Hill-Climbing method and from the GA method with the repair function are compared in Table 3. An initial solution to the Hill-climbing method is generated by the rule-based heuristics as discussed in section 3.1. Multiple runs of 1,000 generation computation are conducted to select the best solution. For the GA method, an initial population is generated at random. The parameters for the GA computation are from the first experiment: population size=500; generation number = 1,000; crossover probability = 0.3; mutation probability = 0.005.

In four experiments, 2, 3, 5 and 9, the GA method generates a solution with a better objective value – a sum of operation time and Δ_{CG}. In the other five experiments, the Hill-Climbing method delivers a better solution. Thus, it seems that the GA method and the Hill-Climbing method show comparable performance in terms of solution quality and computation time. Comparing the two parts of the evaluation function between the GA and the Hill-Climbing solutions show interesting pattern. For the total unloading operation time, solutions from the Hill-Climbing method are better than those from the GA method in all the experiments. On the other hand, the GA method dominates the Hill-Climbing method when we compare the Δ_{CG} part, with only experiment 4 yields a better Δ_{CG} by the Hill-Climbing method.

Table 3. Results from the comparison between the GA with the repair function and the Hill-Climbing

Experiment sets		Hill climbing			GA with the repair function		
Number	**# of containers**	**Evaluation value**	**Operation time (min)**	Δ_{CG}**(m)**	**Evaluation value**	**Operation time (min)**	Δ_{CG} **(m)**
1	**50**	**178.13 17**	**4.53**	**3.6**	**188.73 18**	**8.03**	**0.7**
2	**60**	**206.44 20**	**4.34**	**2.1**	**204.84 20**	**4.34**	**0.5**
3	**95**	**317.99 31**	**2.09**	**5.9**	**314.12 31**	**2.22**	**1.9**
4	**100**	**291.92 28**	**8.42**	**3.5**	**295.29 29**	**0.59**	**4.7**
5	**115**	**436.24 43**	**1.94**	**4.3**	**436.08 43**	**2.38**	**3.7**
6	**120**	**323.09 31**	**3.89**	**9.2**	**355.93 35**	**2.53**	**3.4**
7	**120**	**430.61 42**	**6.31**	**4.3**	**460.10 45**	**8.20**	**1.9**
8	**120**	**476.45 47**	**1.15**	**5.3**	**510.77 50**	**9.17**	**1.6**
9	**120**	**528.44 52**	**4.54**	**3.9**	**526.64 52**	**4.54**	**2.1**

5. DISCUSSION

First, while a larger number of containers to unload generally increase the total unloading operation time, a container stowage configuration – how containers are stacked/located on a container ship – also affects the total unloading operation time. For example, experiments 6-9 have 120 containers to unload, the total operation time for each case varies quite a bit. In a few experiments, an experiment with even a fewer number of containers shows a longer operation time (experiment 3 vs. 4 and 5 vs. 6 in Table 3). Especially, when most containers are stacked in a specific bay of a container ship, the operation time increases because the crane frequently moves along the direction from stem to stern of a MH. While this type of stowage configuration causes a quay crane load balancing problem for a land-based port, it creates an unloading productivity concern for a MH.

Second, the repair function seems to improve the performance of the GA method in this problem. When the repair function is not used and the penalty function for infeasible solution is introduced instead, efficiency of the solution search process decreases. In many cases of the experiments, the GA method without the repair function does not even find a feasible solution and gives a solution with very poor objective value as its final solution. Most likely, it is be because crossover and mutation operations have a high tendency to produce an infeasible solution due to the nature of an order-based GA and the population is quickly filled up with infeasible solutions.

Third, the performance of the GA method does not surpass the Hill-Climbing method. As mentioned in section 4.2, the Hill-Climbing method generates a better solution in 5 of 9 problem instances we tested. The main difference between the Hill-Climbing method and the proposed GA method is the use of crossover operator in the GA method. In the Hill-Climbing method, only mutation operation is included in the search process. In the GA method, on the other hand, crossover operation is a main mechanism with the probability of crossover at 0.3 compared to 0.005 for mutation operation. For our problem, the proposed crossover operation tends to generate many infeasible solutions and much of those are discarded and arbitrarily replaced with feasible solutions by the repair function. We suspect that this process reduces the effectiveness of the search process, thereby negatively affects the quality of final solutions.

Last, the Hill-Climbing method and the GA method shows distinct dominance in the two component of the evaluation function. Solutions from the GA method have a better value for Δ_{CG}, and solutions from the Hill-Climbing are better for the total unloading operation time. We conjecture that this is due to the initial solution used for the Hill-Climbing method along with its search mechanism. A Hill-Climbing method is a local search method, and explores only the vicinity of an initial solution. The initial solution provided to the Hill-Climbing method in our problem is derived from a rule-based method which heavily prefers the operation time minimization. Indeed, the rule-based method used here does not even consider the stability factor. In the previous study by Shin and Lee (2010), the rule-based method is shown to generate a reasonable quality solution in terms of the total operation time. Thus, potential solutions explored in the Hill-Climbing method are expected to have high quality in its operation time measure. In addition, the overall evaluation function used in the experiments, which is a simple sum of the operation time and Δ_{CG}, values the two components equally with $\alpha = \beta$. Since the operation time is two orders of magnitude larger than Δ_{CG}, there seems to be little incentive to aggressively improve in the Δ_{CG} part in the Hill-Climbing search. This is likely to be the reason why there exists clearly distinctive dominance by each method in the two measures.

6. CONCLUSION

A Mobile Harbor presents unique, operational characteristics such as lateral moves by its on-board crane during the unloading process. These lateral moves, from stem to stern or vice versa, negatively affect the unloading productivity. Also, since a MH is a floating structure, container stack arrangement on the deck of a MH and its resultant weight distribution affects the stability of a MH. Thus a container unloading sequence of a MH should be scheduled for both productivity and stability. These differences make this scheduling of MH unloading sequence a unique problem.

This paper presents a mixed integer programming model for the problem, and a heuristic method based on a Genetic Algorithm is proposed. The proposed GA method features a repair function, a crossover operator, and mutation operators designed to fit the specifics of the problem. The effectiveness of the repair function is tested by comparing computational results between the GA method with and without the repair function. It shows that including the repair function in the method dramatically improves the quality of final solutions. The GA method including the repair function is then compared to a Hill-Climbing method. It turns out that the GA method does not outperform the Hill-Climbing method. This is rather unexpected considering the Hill-Climbing method is a local search heuristics. We conjecture that the crossover operator and the selection process used in the current GA method needs improvement.

ACKNOWLEDGEMENTS
This study was supported by the Korean Ministry of Knowledge Economy.

7. REFERENCES

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization, and Machine learning*. Addison Wesley, USA.

Michalewicz, Z. (1999). *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, 3rd, revised and extended edition.

Moccia, L., Cordeau, J.-F., Gaudioso, M. and Laporte, G. (2006). A branch-and-cut algorithm for the quay crane scheduling problem in a container terminal. *Naval Research Logistics, 53(1): 45-59.*

van Kampen, A., Strom, C. and Buydens, L. (1996). Lethalization, penalty and repair functions for constraint handling in the genetic algorithm methodology. *Chemometrics and Intelligent Laboratory Systems, 34(1): 55-68*.

Kim, K. H. and Park, Y. (2004). A Crane Scheduling Method for Port Container Terminals. *European Journal of Operations Research*, 156: 752-768.

Kim, K. Y. and Kim, K. H. (2003). Heuristic Algorithms for Routing Yard-Side Equipment for Minimizing Loading Times in Container Terminals. *Naval Research Logistics, 50(5): 498-514.*

Lee, T., Sung, I., Shin, K. and Nam, H. (2010). Optimal Planning for MH System Operation. *Conference of Society of CAD/CAM Engineers*.

Sammarra, M., Cordeau, J.-F., Laporte, G. and Monaco, M. F. (2007). A tabu search heuristic for the quay crane scheduling problem. *Journal of scheduling, 10(4/5): 327-336.*

Shin, K. and Lee, T. (2010). Container Unloading Scheduling Optimization Problem with MH's Stability Constraint: Near Optimal Solution Searching Method Based on Rule-Based Heuristic and Local Search Method. *Proceedings of Spring Joint Conference of Korean Institute of Industrial Engineers and Korean Operations Research and Management Science Society*.

Whitley, D. (1994). A Genetic Algorithm Tutorial. *Statistics and Computing, 4: 65-85*.

EVALUATING IMPACT OF TRUCK ANNOUNCEMENTS ON CONTAINER STACKING EFFICIENCY

Eelco van Asperen[1*], Bram Borgman[2], and Rommert Dekker[2]

[1]Center for Maritime Economics and Logistics
Erasmus University Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam, The Netherlands
e-mail: vanasperen@ese.eur.nl

[2]Econometric Institute
Erasmus University Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam, The Netherlands
e-mail: rdekker@ese.eur.nl

Abstract: Container stacking rules are an important factor in container terminal efficiency. We build on prior research and use a discrete-event simulation model to evaluate the impact of a truck announcement system on the performance of online container stacking rules. The information that is contained in the announcement, i.e., the expected departure time for an import container, can be used to schedule pre-emptive remarshall moves. These moves can then be performed when the workload is low in order to decrease the export time and the crane workload at peak times.

1. INTRODUCTION

Marine container terminals play a central role in the worldwide distribution of goods that is essential to global supply chains. These container terminals link sea transport via container ships to land transport via trains, barges, and trucks. One of the main problems in container terminals concerns the stacking of containers. Although it is also one of the main advantages of containers, viz. that they can be stacked on top of each other, additional work is required if the bottom container is needed. In that case the top containers have to be moved to another place. These unproductive moves are called reshuffles. They are undesirable as they take additional handling time by the scarce resources.

Accordingly, every terminal needs a stacking strategy. The main objectives of such a strategy are 1) the efficient use of storage space, 2) limiting transportation time from quay to stack and beyond, (and vice versa), and 3) the avoidance of reshuffles. Of course, the importance of each criterion varies from terminal to terminal. Ports like Singapore and Hong Kong have limited land space, so they need efficiently used storage spaces. Note also that these objectives are conflicting: you cannot maximize them all. For example, the third objective would be optimized by having stacks of only one container high; however, this would lead to very inefficient use of storage space and long travel times.

One of the main problems facing a marine container terminal operator is a lack of detailed knowledge regarding the departure time and mode of the containers. While some information may be available at the time of arrival, the mode may change (as containers are rerouted by the shipper) and the departure time is typically a rough estimate for non-scheduled transport modes such as trucks. In this paper we investigate the impact of more accurate departure time information for truck exports, i.e., containers that leave the container terminal by truck.

We investigate a number of scenarios to see whether the length of the period between this information becoming available and the actual departure time can have an influence on the time it takes to remove the container from the stack.

This paper builds on our prior work regarding container terminal stacking (Dekker et al., 2006; van Asperen, 2009; Borgman et al., 2010).. In these studies we focus on the short term decision to allocate an incoming container to a stacking position.

We consider an import container terminal with a high uncertainty regarding the departure time of the containers, such as terminals in the Hamburg–Le Havre range in Europe and the USA. Moreover, we consider online stacking rules, which do not require extensive computations and can be used in many types of stacks and for large numbers of containers. The type of container terminal we consider uses a high degree of automation. This typically implies a tight coupling of the various activities and a reduced flexibility to handle unexpected circumstances. Also, the terminal under consideration has both a high yard utilization and a high berth utilization. This reduces the potential to optimize the storage yard by performing housekeeping moves at times of low utilization. It is therefore relevant to research whether the more accurate departure time information allows for a reduced exit time for containers from the stack.

2. LITERATURE REVIEW

The academic literature on stacking problems is scant but growing, perhaps because the problem does not easily lend itself to analytical solutions (Dekker et al., 2006). In an overview paper on operations research at container terminals, Stahlbock and Voss (2008) looked at a number of aspects of container terminal operations. Among the topics surveyed were stowage planning, berth allocation, crane optimization, terminal transport optimization, and storage and stacking logistics. Their work is an extension of an earlier overview (Steenken et al., 2004), which also contains a paragraph on how stacking is done in practice. For a more detailed overview of marine container terminal stacking algorithms, we refer to Dekker et al. (2006) and Borgman et al. (2010).

The use of Truck Appointment Systems (TAS) for external trucks at container terminals has gained some industry attention but the academic literature is still limited. Some authors refer to “Vehicle Appointment System” or “Container Scheduling System”.

The use of TA systems can be related to solving terminal gate congestion, to improve the turnaround time for external trucks, to reduce the environmental impact of waiting trucks, and to improve stacking operations. The first reported case of a TAS is for the Hong Kong International Terminal (Murty et al, 2005), where it was aimed at making the most of the limited available space in the terminal. The introduction in 2002 of the Lowenthall Bill in California imposed limits of the waiting times for trucks within a terminal in order to reduce the environmental impact. The use of TAS was proposed as one of the possible solutions (Guiliano and O’Brien, 2007). Plata Peredo (2008) describes the use of TAS to minimize terminal handling time for trucks and to reduce the cost for trucking companies. The results of a TAS implementation in Vancouver, Canada were disappointing (Morais and Lord, 2006) as the truck waiting time and congestion were not reduced as expected. The inability of truckers to plan trips was cited as a major cause of this result.

Following a period of heavy congestion for trucks at one of the main container terminals in the port of Rotterdam, a number of studies have looked at the potential of TA systems to reduce this congestion. Van der Heide and van Vliet (2009) give an overview of the IT architecture of a TAS and discuss the potential impact for a non-automated container terminal that uses straddle-carriers. Vlugt (2009) evaluated the potential of a TAS for the automated ECT Delta terminal in the port of Rotterdam. The results showed that the decrease of the expected average truck turnaround time for trucks with an appointment is countered by an increased turnaround time for trucks without an appointment. The net result was that the average turnaround time did not decrease.

So far, the reported results are disappointing with regard to reducing congestion and truck turnaround time. In this paper, however, we focus on the possible impact of a truck appointment system on the performance of the (un)stacking operations in the terminal yard, i.e., we are more interested in the impact on the internal terminal operations. As the focus is on the terminal yard, we are not concerned with gate congestion and environment impacts. We are therefore more interested in the additional information a TAS may generate with regard to the actual time of departure of a stacked container. We will thus use the term “truck announcement” to indicate that the actual coordination between trucking company and terminal operator is outside the scope of this research.

3. BASIC CONCEPTS

We want to investigate the impact of having more information about truck arrivals on container stacking efficiency. As trucks are typically the least controlled transport modes that are served by marine container terminals, having more information should be beneficial. In this paper, we do not use a detailed allocation mechanism for trucks (i.e., the time of arrival is not determined by a slot allocation mechanism). Instead, we consider the situation in which the terminal operator is informed of the arrival time a certain number of hours beforehand.

In our previous research (Dekker et al., 2006; Borgman et al., 2010) we have looked at the performance of a number of container stacking algorithms. The main question we now face is how each stacking rule will benefit from advanced knowledge of truck arrivals. Incoming containers that have to be stored in the stack do not pose problems as long as the information on the departure time and mode are available. The most disruptive influence of truck arrivals occurs when a truck wants to pick-up a container. This container has to be retrieved ("unstacked") from the yard, potentially causing one or more unproductive moves (reshuffles) in order to access the container. With advance information on the precise time and mode of departure we may be able to perform the reshuffle operations using idle time of the automated stacking crane (ASC) between the moment the departure time information becomes available and the actual departure time. While this in itself will not reduce the number of unproductive moves, we may be able to perform these unproductive moves at a time when they do not impact the performance of the system, i.e., during idle time of the automated stacking crane.

If we consider a single lane of the type of container terminal under investigation, we see that there is a single rail-mounted stacking crane that has to perform all the stacking moves for that lane. In this front-end interchange design there is an interchange point between the horizontal transport (AGV, straddle carrier or yard truck) and vertical transport (by the automated stacking crane) at either end of the lane. We can distinguish between containers that are moving into the lane (i.e., that are being stacked) and containers that are moving out of the lane. Containers can enter and leave the lane at two sides: at the quay side (for containers that are coming from or going to deep-sea ships) and at the land side (for all other modes of transport). Figure 1 provides a schematic overview of the terminal layout. The layout of this terminal has the stacking lanes perpendicular to the quay. Each lane has a length, a width, and a height.

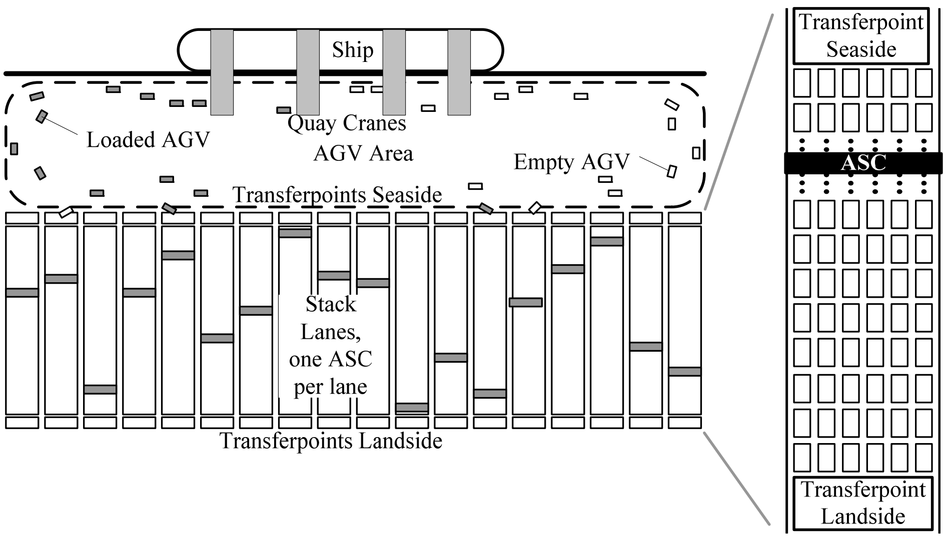


Figure 1. Terminal Layout

The overall approach of the experiments in this paper is focused on the operational decisions that have to be made by terminal operators. Specifically, we take the arrivals and departures that are specified as part of the generator output (v.i.) and perform these operations. There is no global optimization or explicit planning; the operations are performed one at a time, i.e. in a greedy fashion, whenever a container arrives or has to leave the stack. We do not consider future events such as other incoming containers. The prearrival information with regard to the truck arrivals is used as follows:

1. The container is locked, i.e. no container may be placed on top of it;
2. If the ASC has an empty work queue, then it checks whether the locked container is on top of the pile. If it is not, the ASC tries to reshuffle the containers on top of the locked container in order to free it for easy retrieval.

In other words, if the crane is going to be idle between the moment we receive the announcement and the arrival time of the truck, we may be able to perform the required reshuffles without interfering with other jobs for the crane.

4. SIMULATION MODEL

The simulation model that was developed for the experiments in this and previous papers consists of two major components: a generator and a simulator. Both programmes have been implemented in the Java programming language; the simulator uses the SSJ discrete-event simulation library (L'Ecuyer and Buist, 2005).

The generator program creates arrival and departure times of the containers. The generator is based on the same data as the generator in Dekker et al. (2006), including sailing schedules and a modal-split matrix. For the experiments in this paper, the modal-split matrix has been updated to reflect the recent increase in truck volume in comparison with the scenarios studies in Dekker et al (2006). The output of the generator is a file that contains the ship arrivals, details of the containers to be discharged and loaded, and the specification of the destination of each container. The departure time is specified as the planned (a.k.a. expected) departure time and the actual (a.k.a. real) departure time. The destination can be another deep-sea vessel or (for import containers) a short-sea vessel, barge, train or truck. For each container the location of the individual container within a ship is specified. The generator takes the detailed quay crane sequences for loading and discharging into account.

The simulator program reads the output of the generator and performs the stacking algorithms. The core of the simulator itself is deterministic: the stochastic components are in the generator and, optionally, in the stacking algorithm. This setup facilitates a comparison of stacking algorithms as any changes in the statistical output of the simulator must be caused by the stacking algorithm. This approach can be considered a way to implement common random numbers. More details of the generator and simulator are described in Borgman (2009).

5. EXPERIMENTAL SETUP

The experiments in this paper all use the following configuration. Experiments are run for a 15 week period, of which three weeks are used for warm-up (to initialize the stack). We have used the generator program, discussed in section \ref{model}, to generate an arrivals file containing 175,360 arriving containers during the 12 weeks of simulation (45.2\% 20ft, 48.4\% 40ft, and 6.4\% 45ft). Some stacking rules have a stochastic component (such as selecting a position at random). As in our previous experiments we have used ten replications to get statistically robust results. These replications are used to compute the 95\%-confidence intervals of the means.

We assume that there are sufficient AGVs and straddle carriers to ensure that these resources do not act as bottlenecks. The basic configuration for the stacking area was adapted from the earlier work by the same authors (Borgman et al., 2010), but resized to accommodate the increased number of containers and the various container sizes. The stack configuration is now 19 lanes, each consisting of 6 segments, which in turn have 55 positions of 20 ft. The length of the lanes was increased because the average size of a container and the average dwell time have increased. Keeping the lanes short would thus result in less containers per lane, and hence less work for the ASCs. The increase in the number of lanes was still required to accommodate the increase in the absolute number of containers. Each lane still has a single ASC per lane, and the maximum stacking height is still three containers. The average residence time of a container is 4.6 days and the average utilization of the yard is 67%.

6. EXPERIMENTS

In this section we will present our experiments with a number of stacking rules. Before we discuss the stacking rules we should first consider how we are going to evaluate the stacking rules. We measure the performance of a stacking algorithm with the following statistics:

Exit Time (ETQ and ETL) The exit time is the time (in hours) it takes to remove a container from the stack and have it ready for onward transport (to the quay or to a truck/train/barge). This time is measured for each side (quay-side and land-side) of the stack and will be listed as ETQ and ETL respectively. The exit time is the main performance indicator for a stacking algorithm. It is negatively influenced by stacking further away from the exit point and by any reshuffles that are needed to retrieve the container. When a container enters the stack, the time it takes to perform this operation is determined by the workload of the ASC (how many jobs are in the current job queue) and the time it takes the ASC to move the container to its position. There are no reshuffles when containers are stored in the stack; reshuffles only occur when a container has to leave the stack. As we consider these two statistics the main performance indicators, we also report the 90% percentile of the mean.

ASC Workload (ACS) The automated stacking cranes are critical components for the overall performance so we measure the percentage of time that the ASC's are busy. (The ASC workload will be denoted as ASC in the results.)
Reshuffles (RDC and ROC) For the unproductive reshuffle moves, we measure the number of reshuffles (denoted as RDC) as a percentage of the total number of container movements. To get an indication of the number of reshuffles that happen per move, we also measure the reshuffle occasions (as a percentage of the total number of container movements, denoted as ROC); a single reshuffle occasion implies one or more reshuffles. These numbers are not absolute indicators of performance as the time of the reshuffle is not taken into consideration. A reshuffle that occurs when the workload is low has less impact on the overall performance than a reshuffle during a peak workload, for example when (un)loading a very large vessel.

The stacking algorithms are described in more in detail in Dekker et al. (2006) and Borgman et al. (2010). A short description for each of the algorithms is listed in table 1. We first test the impact of truck announcement with the reference algorithms (random stacking and leveling). To test the potential benefit of having this more accurate information on the container departure time, we have run both algorithms without the truck announcement and with announcement a number of hours before the departure time (ranging from 0.5 to 24 hours). The results are listed in table 2.

Table 1. Stacking Algorithms

RS	Random Stacking: The new container is placed at a randomly chosen allowed location, with every allowed location having an equal probability of being chosen. This algorithm is also applied for reshuffling, with the difference that we then only want to search the lane the container is in.
LEV	Leveling: The idea is to fill lanes in layers, so that all empty ground positions are filled with containers first, before containers are stacked upon others. The stacking lane is filled from the transfer point quayside on.
RS-DT	The Random-Stacking with Departure Times algorithm searches for a random pile where the top container's departure time is after the new container's departure time.
RS-DTC	This is the same algorithm as RS-DT but with a small number of classes of departure times.
LDT	The Leveling with Departure Times algorithm modifies the basic leveling algorithm to only stack a container on top of a container with a later expected departure time.
LDT-DTC	Leveling with Departure Time Classes uses a small number of departure time classes rather than the actual departure time.
TPRL	The Transfer Point Random Level algorithm picks a level (height) at random and then looks for a location that is closest to the transfer point.
TVR	The Traveling Distance versus Reshuffling algorithm balances the two components.
TVR-DTC	With this algorithm we modify the TVR algorithm to use a small number of classes rather than the actual departure time.
TVR-DTC-MD	This modification of the TVR-DTC algorithm minimizes the difference in the departure time class.

In table 2 we have listed the results of the two reference algorithms, random stacking and leveling. The statistically significant results for the TAS (in comparison with the results without the TAS, at the 95% level) for the main performance indicators, ETQ and ETL, have been marked with an asterisk. We can see that the truck announcement has an impact on the exit times for the random stacking algorithm. The small differences between the exit times beyond an announcement time of four hours suggest that even an announcement just before the pick-up time is useful. For the leveling algorithm, the impact of the shorter announcements (0.5 and 1 hour) is not significant. The longer announcement does result in a decrease in the exit time for both the land and the sea side. There appears to be little value in announcements longer than twelve hours in advance.

Table 2. Results for the Reference Stacking Algorithms.

Stacking algorithm	TAS (hrs)	ROC (%)	RDC (%)	ASC (%)	ETQ (hrs)	ETL (hrs)	90% ETQ (hrs)	90% ETL (hrs)
RS	0	65.11	97.80	53.81	0.39	0.33	0.91	0.75
RS	0.5	65.11	97.81	53.92	0.38	0.29*	0.88*	0.72
RS	1	65.11	97.71	53.91	0.37*	0.28*	0.85*	0.69*
RS	4	65.13	97.71	53.93	0.35*	0.26*	0.78*	0.61*
RS	12	65.33	97.87	53.97	0.34*	0.25*	0.76*	0.58*
RS	24	65.69	98.23	54.05	0.35*	0.26*	0.78*	0.59*
LEV	0	56.70	66.73	50.69	0.50	0.43	1.27	1.11
LEV	0.5	56.71	66.59	50.00	0.37	0.30	0.85	0.69
LEV	1	56.66	66.57	50.21	0.40	0.33	0.94	0.78
LEV	4	56.72	66.72	49.94	0.35*	0.29*	0.76*	0.59*
LEV	12	56.96	67.26	49.87	0.33*	0.26*	0.71*	0.55*
LEV	24	57.49	68.47	49.98	0.32*	0.26*	0.69*	0.53*

In table 3, we show the results of all our other stacking algorithms, tested with no TAS and a TAs announcement time of 12 hours, which proved to have the best performance in most cases. The mixed and unmixed variants of the algorithms indicate whether transshipment (or sea-sea) containers are mixed in the same piles as import or export containers; 'real' and 'exp' indicate the use of the actual and the expected time of departure in the stacking algorithm, respectively.

From the table, we can see that for most algorithms, a TAS does not significantly improve performance, even though in almost all cases the ETL is slightly lower. The exceptions are RS, LEV (both as previously noted), RS-DT, RS-DTC and TPRL. These algorithms are also the ones that performed worst without the use of a TAS. This indicates that TAS is a potentially useful feature. The reason that the poorer algorithms have improved might lie in the fact that the TAS strategy (in particular the locks it sets) prohibits these algorithms from making bad decisions. The better algorithms already make few mistakes and hence leave little room for improvement.

Interestingly, we do not find that the lower ETL adversely affects the exit times for other containers (ETQ), which was reported by Vlugt (2009). We do however note that our TAs model is significantly different and the two may not be comparable.

Table 3: Result of the Experiments for All Stacking Algorithms.
Significant differences between experiments are indicated by a *-sign (at the 95% level).

Stacking algorithm	TAS	ROC	RDC	ASC	ETQ (hrs)	ETL (hrs)	90% ETQ (hrs)	90% ETL (hrs)
LEV	0	11.58	13.37	42.48	0.50	0.43	1.27	1.11
LEV	12	56.96	67.26	49.87	0.33*	0.26*	0.71*	0.55*
RS	0	11.58	13.37	42.48	0.39	0.33	0.91	0.75
RS	12	65.33	97.87	53.97	0.34*	0.25*	0.76*	0.58*
LDT (mixed, exp)	0	11.58	13.37	42.48	0.24	0.23	0.47	0.41
LDT (mixed, exp)	12	14.73	16.78	42.94	0.24	0.22	0.47	0.39
LDT (unmixed, exp)	0	12.95	15.53	42.31	0.21	0.20	0.36	0.32
LDT (unmixed, exp)	12	15.28	18.19	43.29	0.20	0.19	0.35	0.30
LDT (mixed, real)	0	2.86	3.66	41.53	0.20	0.19	0.33	0.30
LDT (mixed, real)	12	6.54	7.80	42.07	0.20	0.19	0.34	0.30
LDT (unmixed, real)	0	5.46	6.88	41.85	0.18	0.17	0.30	0.26

LDT (unmixed, real)	12	8.84 10	.80 42	.92 0.	18	0.17	0.30	0.26
RS-DT (exp)	0 26	.73 38	.63 45	.95 0.	23	0.21	0.45	0.38
RS-DT (exp)	12 2	8.04 3	9.76 4	6.16 0	.22* 0	.19* 0	.43* 0	.34*
RS-DT (real)	0 23	.18 34	.50 45	.57 0.	21	0.20	0.41	0.34
RS-DT (real)	12 2	5.06 3	6.65 4	5.90 0	.21* 0	.18* 0	.39* 0	.31*
LDT-DTC (exp)	0 27	.91 34	.12 45	.42 0.	29	0.25	0.60	0.49
LDT-DTC (exp)	12 30	.43 37	.39 46	.77 0.	28	0.23	0.57	0.43
LDT-DTC (real)	0 28	.62 34	.84 45	.42 0.	29	0.25	0.61	0.49
LDT-DTC (real)	12 31	.08 38	.19 46	.77 0.	28	0.24	0.59	0.46
RS-DTC (exp)	0 54	.73 78	.96 51	.04 0.	34	0.29	0.80	0.65
RS-DTC (exp)	12 5	4.91 7	9.09 5	1.14 0	.31* 0	.23* 0	.67* 0	.50*
RS-DTC (real)	0 55	.09 79	.44 51	.05 0.	34	0.29	0.80	0.65
RS-DTC (real)	12 5	5.15 7	9.39 5	1.15 0	.31* 0	.23* 0	.68* 0	.50*
TPRL	0 55	.27 83	.92 47	.48 0.	29	0.25	0.56	0.45
TPRL	12 55	.74 84	.28 47	.78 0.	26	0.21* 0.	50	0.39
TVR (p=0)	0 54	.57 84	.47 46	.38 0.	27	0.23	0.51	0.41
TVR (p=0)	12 55	.01 84	.74 46	.71 0.	26	0.21	0.49	0.37
TVR (p=0.03)	0 54	.65 76	.99 46	.71 0.	26	0.23	0.48	0.38
TVR (p=0.03)	12 55	.10 77	.46 46	.98 0.	25	0.20	0.46	0.35
TVR-DTC-MD (p=0)	0 33	.42 46	.99 43	.23 0.	21	0.18	0.37	0.30
TVR-DTC-MD (p=0)	12 34	.23 47	.87 43	.42 0.	23	0.19	0.40	0.30
TVR-DTC-MD (p=0.03)	0 15	.58 20	.97 42	.09 0.	18	0.16	0.29	0.26
TVR-DTC-MD (p=0.03)	12 16	.66 22	.16 42	.27 0.	18	0.17	0.30	0.26
TVR-DTC (p=0)	0 29	.97 43	.40 42	.67 0.	21	0.19	0.36	0.30
TVR-DTC (p=0)	12 30	.85 44	.31 42	.91 0.	20	0.17	0.34	0.28
TVR-DTC (p=0.03)	0 19	.28 26	.88 42	.52 0.	19	0.17	0.32	0.27
TVR-DTC (p=0.03)	12 20	.30 28	.08 42	.71 0.	19	0.17	0.33	0.27

7. CONCLUSION

In this paper, we experimented with a Truck Announcement System to try to improve container stacking efficiency. For several strategies, we found that an announcement time of 0.5 to 24 hours significantly improves the efficiency, and it matters little how far in advance the warning is given. However, this only applies to several basic strategies. More elaborate strategies, using more information about the containers, did not significantly benefit from the announcements.

From these experiments we also conclude that further research is needed to evaluate the value of truck announcements. A number of results proved to be statistically non significant. It is however striking that all stacking algorithms show a decrease in the exit times. Even though the individual experiments may not show a statistically significant benefit, we suspect that there may indeed be a benefit and that further research is required.

REFERENCES

Asperen, Eelco van (2009). Essays on Port, Container, and Bulk Chemical Logistics Optimization. PhD Thesis, Erasmus Research Institute In Management, Erasmus University Rotterdam. ISBN 978-90-5892-222-9.

Borgman, Bram (2009). Improving container stacking efficiency using simulation. Master's thesis, Erasmus University Rotterdam, Erasmus School of Economics. URL http://hdl.handle.net/2105/5040. Economics & Informatics programme.

Borgman, Bram, Eelco van Asperen, and Rommert Dekker (2010). Online rules for container stacking. OR Spectrum, 32(3):687–716. doi: 10.1007/s00291-010-0205-4. URL http://dx.doi.org/10.1007/s00291-010-0205-4. Open Access.

Dekker, Rommert, Patrick Voogd, and Eelco van Asperen (2006). Advanced methods for container stacking. OR Spectrum, 28:563–586.

Duinkerken, Mark B., Joseph J.M. Evers, and Jaap A. Ottjes (2001). A simulation model for integrating quay transport and stacking policies in automated terminals. In Proceedings of the 15th European Simulation Multiconference (ESM2001), Prague, 2001. SCS.

Guiliano, G. and T. O'Brien (2007). Reducing port-related truck emissions: The terminal gate appointment system at the ports of Los Angeles and Long Beach. Transportation Research Part D, 12(7):460–473.

L'Ecuyer, Pierre and Eric Buist (2005). Simulation in Java with SSJ. In M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, editors, Proceedings of the 2005 Winter Simulation Conference, pages 611–620.

Morais, Philipppe and Elisabeth Lord (2006). Terminal appointment system study. TP14570E.

Murty, K.G., Y-w. Wan, J. Liu, M.M. Tseng, E.L. Leung, K-k. Lai, and H.W.C. Chiu (2005). Hongkong international terminals gains elastic capacity using a data-intensive decision-support system. Interfaces, 35(1):61–75.

Plata Peredo, I.E. (2008). Fluidity at the terminal gate: Truck appointment systems as remedy. Master's thesis, Maritime Economics and Logistics, Erasmus University Rotterdam, School of Economics.

Stahlbock, Robert and Stefan Voss (2008). Operations research at container terminals: a literature update. OR Spectrum, 30:1–52.

Steenken, Dirk, Stefan Voss, and Robert Stahlbock (2004). Container terminal operation and operations research a classification and literature review. OR Spectrum, 26:3–49.

van der Heide, S. and H. van Vliet (2009) Truck appointment systems, Seminar Report Economics and ICT, Erasmus University Rotterdam.

Vlugt, D. (2009). Truck congestion at and in the terminal; a simulation study. Bachelor thesis Econometrics in Logistics, Erasmus University Rotterdam.

A STUDY OF SCHEDULING ALGORITHMS FOR MOBILE HARBOR WITH AN EXTENDED M-PARALLEL MACHINE PROBLEM

Inkyung Sung[1], Hochang Nam[1], and Taesik Lee[1*]

[1]Department of Industrial and Systems Engineering
Korea Advanced Institute of Science and Technology
335 Gwahangno, Yuseong-Gu
Daejeon, Republic of Korea
E-mail: taesik.lee@kaist.edu

Abstract: Mobile Harbor is a ship, or a movable floating structure with container loading/unloading equipment on board, so that it can work with a container ship anchored in sea. In some sense, a Mobile Harbor is equivalent to a berth with a quay crane of a conventional port. For a Mobile Harbor-based system to handle a large volume of containers, multiple units of Mobile Harbor are required. As such, operation schedule for the units is essential to the productivity of the system. In this paper, a method to compute a scheduling solution for Mobile Harbor is proposed. It determines the operation and time sequence of jobs by the units to minimize the makespan of container handling operation. This problem is formulated with Mixed Integer Programming (MIP), based on an m-parallel machine problem. In addition, a heuristic approach using a Genetic Algorithm is developed to obtain a near optimal solution with reduced computation time.

1. INTRODUCTION

Volume of maritime container transportation has increased steadily. The continuous increase in the container volume has stimulated the building of very large container ships that is capable of carrying over 8,000TEU[1], in an effort to reduce the transportation costs. With the introduction of such large container ships, container terminals are now facing challenges to dramatically improve their service capability to effectively serve the container ships. The challenges include providing sufficient water depth at their berths and in their sea routes to accommodate a very large container ship, and improving container handling productivity to reduce port staying time for container ships. Resolving these problems with conventional approaches – expanding existing ports or building new ones – requires massive investment and sometimes causes environmental concerns.

Mobile Harbor has been introduced by KAIST as an alternative solution to this problem. Mobile Harbor is a container transportation system that can load/unload containers from an anchored container ship in the open sea. It can transfer containers from a container ship to a container terminal, and vice versa. As container ships do not need to directly call at the container terminal, there is no water level requirement for a port. Mobile Harbor also makes it possible to transfer containers to locations where container handling infrastructure does not exist. The concept of Mobile Harbor is shown in Figure 1.

Figure 1. Concept of Mobile Harbor

Initial cost and operating cost of a Mobile Harbor system are significant, and thus, increasing productivity of Mobile Harbor operation is a key to improving its practical feasibility. With a given number of Mobile Harbor units and related resources, we would want to maximize service performance. This can be achieved by developing an optimal scheduling of Mobile Harbor operation. This paper proposes scheduling methods for the operation of a fleet of Mobile Harbor units. Specifically, we study the problem of allocating Mobile Harbor units to a series of incoming container ships and determining detailed work schedule for each Mobile Harbor unit.

[1] TEU stands for twenty-foot equivalent unit. It is a measure of 20 ft container

Scheduling Mobile Harbor units is similar to quay crane scheduling problems. A quay crane scheduling problem is typically modeled as an m-parallel machine problem. Due to its similarity to a quay crane scheduling problem, the Mobile Harbor scheduling is formulated in this study as a modified m-parallel machine problem. M-parallel machine problems formulated for typical quay crane scheduling are known to be difficult to solve for exact solutions. Thus, many of the quay crane scheduling research propose an efficient method to overcome the computational complexity of exact solution approaches. Kim *et al.* (2004) provided a greedy randomized adaptive search procedure to obtain a near-optimal solution. Moccia *et al.* (2006) proposed Branch and Cut algorithm to solve the quay crane scheduling on the small and medium size problems. Sammarra *et al.* (2007) proposed a tabu search heuristic algorithm. Lee *et al.* (2008) provided a genetic algorithm. In this paper, a basic m-parallel machine problem has been modified to incorporate the unique characteristics of Mobile Harbor operation. As with many quay crane scheduling researches, we also propose a heuristic method.

This paper is organized as follows. In Section 2, operational characteristics of a Mobile Harbor system are clarified to define the Mobile Harbor scheduling problem. A mathematical model representing a Mobile-Harbor-based port system is formulated by Mixed Integer Programming (MIP) in Section 3. Section 4 presents a heuristic algorithm using a Genetic algorithm (GA) to obtain a near-optimal solution for the problem. Performances of the MIP model and GA-based heuristic algorithm are compared in Section 5. Discussion and Concluding remarks are presented in Section 6 and 7, respectively.

2. MOBILE HARBOR SCHEDULING PROBLEM

In this paper, Mobile Harbor scheduling refers to the allocation of multiple units of Mobile Harbor to a multiple number of incoming container ships so that container ships' staying time is minimized. We only consider unloading of containers from container ships to a land-side berth. Loading part of the operation has been excluded in the modeling of this paper.

There are a few characteristics unique to Mobile Harbor that need to be considered in the scheduling problem. First, a Mobile Harbor unit has a finite container carrying capacity (e.g. 250 TEU). Once its maximum capacity is reached, it must visit a land-side berth to unload its containers to be available for a next service. This time – travel to a land-side berth, unload at the berth, and travel back to a container ship – is considered as a downtime of a Mobile Harbor unit, from the scheduling perspective. Second, a physical dimension of a Mobile Harbor unit has an implication for its scheduling. In particular, Mobile Harbor's length – stem to stern – defines its work space. A container unloading system, i.e. container crane, can move only within a range constrained by the Mobile Harbor unit's length dimension. In case of 250 TEU model, this length dimension is approximately 70 meters, and the onboard crane can travel for about 50 meters from aft to bow (equivalent to four 40-ft bays long). At one docking position, a Mobile Harbor unit can work with containers located within its working range. For 250 TEU Mobile Harbor, it can unload containers on a container ship within four 40-ft bay range at a single docking position. Third, multiple numbers of Mobile Harbor units can work on a container ship concurrently if docking positions for each Mobile Harbor unit can be properly assigned for a concurrent operation. A Mobile Harbor unit docks with a container ship at any position on either side (port or starboard) of a container ship unless there is a Mobile Harbor unit working nearby within a desired separation distance. When two Mobile Harbor units dock on the opposite side of a container ship, their separation distance for safety is shorter than when they dock on the same side. See Figure 2.

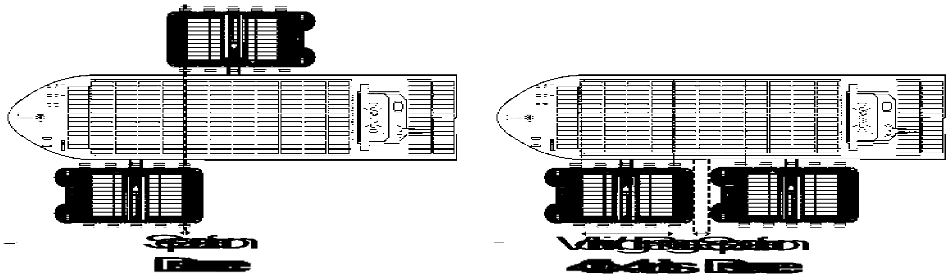


Figure 2. Mobile Harbor – container ship

The first step in the Mobile Harbor scheduling problem is to define jobs. Containers to be unloaded from a container ship are divided into a finite number of jobs. A job is defined as a set of containers to be unloaded from a container ship by a single Mobile Harbor unit. Maximum number of containers of a single job should be less than the container carrying capacity of one unit of the Mobile Harbor. Also, containers that belong to a job must be located within a Mobile Harbor unit's working range – four consecutive bays on a container ship in case of 250TEU model –

due to a work range constraint of a Mobile Harbor. For a given number of containers to be unloaded and their stowage plan, many different set of jobs can be defined. Figure 3 shows an example of two feasible job definitions for 380 containers with the same stowage plan. Ideally, job definition must be solved as a part of the whole scheduling optimization, as a choice of job set will affect the scheduling solution. In this study, we simply define jobs such that a minimum number of jobs are obtained.

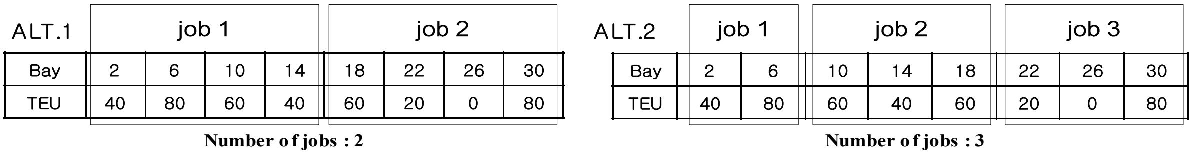


Figure 3. Example of a job set definition. A job set with a smaller number of jobs (e.g. ALT 1) is used as an input to the scheduling problem.

With a set of jobs defined, a Mobile Harbor scheduling problem is to assign each job to a set of available Mobile Harbor units, determine its docking side (port or starboard), and determine each job's starting time, under the following assumptions:

- a Mobile Harbor unit has a capacity limit (e.g. 250 TEU),
- it has a downtime once its maximum carrying capacity is reached,
- it docks at any position and side of a container ship,
- at one docking position, it can handle containers on a container ship within its working range (e.g. for 250 TEU Mobile Harbor unit, its working range is four 40-ft bay long),
- when docking, there should be a minimum separation distance between adjacent Mobile Harbor units, for safety or work interference reason,
- minimum separation distance is different for the same-side docking and the opposite-side docking, and
- container ships arrive according to their calling schedule. This means that not all jobs are available at the beginning of a scheduling period, and each job's available time is determined by a container ship's arrival time.

While the concurrent operation of multiple Mobile Harbor units is allowed in principle, some of the jobs cannot be operated concurrently. In other words, for some pair of jobs, one job must be completed first before the other job can start. There are two conditions that prevent concurrent operations of jobs: conflict condition and semi-conflict condition. A conflict condition occurs when two (or more) jobs have containers stored in a common, overlapping bay on a container ship, as shown in Figure 4(a). This happens when a large number of containers are located in a few adjacent bays on a ship so that a single Mobile Harbor unit cannot clear all containers in the area. On the other hand, when two jobs are distanced away less than a separation distance (same-side docking), these two jobs are under a semi-conflict condition (Figure 4(b)). The term "semi-" is added to indicate that these two jobs can be conducted concurrently if two Mobile Harbor units dock at an opposite side from each other.

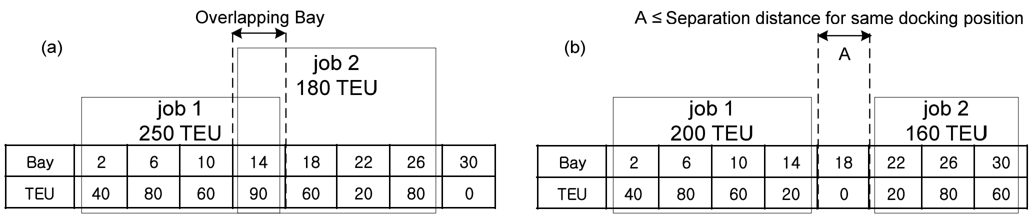


Figure 4. Conflict condition and Semi- conflict
(a) Conflict condition for jobs: Job1 and Job 2 must be executed sequentially regardless of choices of docking side,
(b) Semi- conflict condition for jobs: Job 1 and Job 2 can be executed concurrently if operated from the opposite sides. If same side docking, they must be done sequentially.

3. DESCRIPTION OF MATHEMATICAL MODEL

A mathematical model formulated by MIP is represented in this section. The model builds upon a quay crane scheduling problem by Kim *et al*. (2004). To incorporate the operational characteristics of Mobile Harbor, the original model has been modified significantly. The notations used in the MIP model are shown below:

Index

i, j, u	Job ID
k	Mobile Harbor ID
v	Docking position of Mobile Harbor on a container ship (port, starboard)
c	Container ship ID

Problem data

n_i	The number of containers in job *i*
p_i	The processing (unloading) time of job *i*
a_c	The arrival time of container ship *c*
M	A sufficiently large number
tr	The travel time (one-way) of a Mobile Harbor between a container ship and a land-side berth
teu	The time to unload 1 TEU [min/TEU]
$MH_{capacity}$	The container carrying capacity of one unit of the Mobile Harbor

Index set

Ω	The set of all jobs
Φ	The set of pairs of jobs under a conflict condition. These pairs of jobs cannot be operated concurrently
C	The set of container ships
K	The set of Mobile Harbor units
Θ	The set of pairs of jobs under a semi- conflict condition. These pairs of jobs cannot be operated concurrently from the same side, but can be if from the opposite side
Ψ_c	The set of jobs of container ship *c*

Decision variables

X_{ij}^k	1, if Mobile Harbor *k* executes job *j* after completing job *i*, without any other job in between. 0, otherwise
S_{iv}^k	1, if Mobile Harbor *k* execute job *i* at side *v* of a container ship. 0, otherwise
Y_c	The completion time of container ship *c*; This is the time when all the jobs of container ship *c* are completed
D_i	The completion time of job *i*
Z_{ij}	1, if job *j* starts after the completion of job *i*. 0, otherwise
$Cont_{ij}^k$	1, if Mobile Harbor *k* executes job *j* after completing job *i* without visiting a land-side berth. 0, otherwise
L_{ij}^k	For a Mobile Harbor *k* that has completed job *i* - a predecessor to job *j*, the total number of containers currently stored on Mobile Harbor *k* after finishing job *j*. When Mobile Harbor *k* executes job *j* immediately after job *i* ($Cont_{ij}^k$ = 1), this is a sum of the number of containers from job *i* and *j*.
t_{ij}^k	The downtime of Mobile Harbor *k* between job *i* and job *j*. If $Cont_{ij}^k = 1$, t_{ij}^k is zero.

The MIP model

$$Minimize \sum_c Y_c \tag{1}$$

subject to

$$\sum_{j\in\Omega\cup\{T\}} X_{0j}^k = 1 \qquad \forall k \in K, \tag{2}$$

$$\sum_{i\in\Omega\cup\{0\}} X_{iT}^k = 1 \qquad \forall k \in K, \tag{3}$$

$$\sum_{k\in K}\sum_{i\in\Omega\cup\{0\}} X_{ij}^k = 1 \qquad \forall j \in \Omega, \tag{4}$$

$$\sum_{j\in\Omega\cup\{T\}} X_{ij}^k - \sum_{j\in\Omega\cup\{0\}} X_{ji}^k = 0 \qquad \forall i \in \Omega, \forall k \in K, \tag{5}$$

$$D_i - p_i \ge a_c \qquad \forall i \in \Psi_c, \forall c \in C, \tag{6}$$

$$Y_c \ge D_i \qquad \forall i \in \Psi_c, \forall c \in C, \tag{7}$$

$$D_i + p_j - D_j \le M(1 - Z_{ij}) \qquad \forall i, j \in \Omega, \tag{8}$$

$$D_j - p_j - D_i \le M \cdot Z_{ij} \qquad \forall i, j \in \Omega, \tag{9}$$

$$Z_{ij} + Z_{ji} = 1 \qquad \forall (i,j) \in \Phi, \tag{10}$$

$$\sum_{k\in K}\sum_{v\in V} S_{iv}^k = 1 \qquad \forall i \in \Omega, \tag{11}$$

$$\sum_{i\in\Omega\cup\{0\}} X_{ij}^k - \sum_{v\in V} S_{jv}^k = 0 \qquad \forall j \in \Omega, \forall k \in K, \tag{12}$$

$$\sum_{k\in K} S_{i1}^k - \sum_{k\in K} S_{j2}^k \le M(Z_{ij} + Z_{ji}) \qquad \forall (i,j) \in \Theta, \tag{13}$$

$$\sum_{k\in K} S_{i2}^k - \sum_{k\in K} S_{j1}^k \le M(Z_{ij} + Z_{ji}) \qquad \forall (i,j) \in \Theta, \tag{14}$$

$$X_{ij}^k \ge Cont_{ij}^k \qquad \forall i, j \in \Omega, k \in K, \tag{15}$$

$$n_j - L_{ij}^k \le M \cdot (1 - X_{ij}^k) \qquad \forall i \in \Omega \cup \{0\}, \forall j \in \Omega, k \in K, \tag{16}$$

$$L_{ij}^k \le M \cdot X_{ij}^k \qquad \forall i \in \Omega \cup \{0\}, \forall j \in \Omega, k \in K, \tag{17}$$

$$\sum_{u\in\Omega\cup\{0\}} L_{ui}^k + n_j - MH_{capacity} \le M(1 - Cont_{ij}^k) \qquad \forall i, j \in \Omega, \forall k \in K, \tag{18}$$

$$\sum_{u\in\Omega\cup\{0\}} L_{ui}^k + n_j - L_{ij}^k \le M(1 - Cont_{ij}^k) \qquad \forall i, j \in \Omega, \forall k \in K, \tag{19}$$

$$t_{ij}^k \le M(1 - Cont_{ij}^k) \qquad \forall i, j \in \Omega, \forall k \in K, \tag{20}$$

$$2 \cdot tr + (teu \cdot \sum_{u\in\Omega\cup\{0\}} L_{ui}^k) - t_{ij}^k \le M \cdot Cont_{ij}^k \qquad \forall i, j \in \Omega, \forall k \in K, \tag{21}$$

$$D_i + t_{ij}^k + p_j - D_j \le M \cdot (1 - X_{ij}^k) \qquad \forall i \in \Omega \cup \{0\}, \forall j \in \Omega, k \in K, \tag{22}$$

$$Y_c \ge 0 \qquad \forall c \in C, \tag{23}$$

$$D_i \ge 0 \qquad \forall i \in \Omega, \tag{24}$$

$$L_{ij}^k \ge 0, t_{ij}^k \ge 0 \qquad \forall i \in \Omega \cup \{0\}, \forall j \in \Omega, k \in K, \tag{25}$$

$$X_{ij}^{k} = 0\, or\, 1, \qquad \forall i \in \Omega \cup \{0\}, \forall j \in \Omega \cup \{T\}, \forall k \in K \qquad (26)$$

$$Z_{ij}, S_{iv}^{k}, Cont_{ij}^{k} = 0\, or\, 1 \qquad \forall i, j \in \Omega, \forall k \in K \qquad (27)$$

An objective function (1) minimizes the sum of the completion times of all container ships. Constraints (2) and (3) set the first job and last job, respectively, of each Mobile Harbor. Job *0* and *T* are dummy jobs. Constraint (4) guarantees that all jobs are assigned to one of the Mobile Harbor units. Constraint (5) ensures the continuity of a job sequence assigned to a Mobile Harbor. Constraint (6) requires all jobs completed some time after the arrival time of the corresponding container ship. Constraint (7) determines the completion time of each container ship. Constraints (8)~(10) guarantee that two jobs in the conflict condition are executed sequentially. Constraints (11)~(14) determine concurrent or sequential execution of a job pair in the semi-conflict condition and their docking position (port/starboard). Constraints (15)~(22) allow a Mobile Harbor unit to take on a job immediately after finishing the previous job without visiting a land-side berth, if the unit's remaining capacity is larger than the number of containers of the next job.

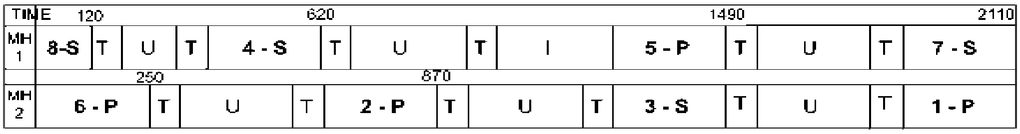


Figure 5. Example of an optimal solution for Mobile Harbor scheduling using the MIP model
#: job ID, P(Port)/S(Starboard): Docking position of Mobile Harbor for a job, T: Travel time between the container ship and a land-side berth, U: Unloading time for containers to a land-side berth, I: Idle time

Figure 5 is an example of an optimal solution for Mobile Harbor scheduling from the MIP model. The MIP model required 14 hours to obtain an optimal solution for the above example. The MIP model was implemented using CPLEX and the experiment was conducted on a Core2 Quad CPU 2.40GHz with 3GB of RAM.

4. GENETIC ALGORITHM

The MIP model presented in section 3 requires a considerable amount of computation time. It is expected that solving the MIP model for a problem with a large number of jobs is practically infeasible. To overcome this problem, a heuristic method based on a Genetic Algorithm (GA) is proposed to obtain a near-optimal solution within a reasonable amount of computation time.

4.1 Solution representation

A sequence of jobs is represented by a chromosome. In a chromosome, the position where a job is located indicates the identification number of the Mobile Harbor unit each job is assigned to. When the total number of Mobile Harbor units in a system is *m*, a job at the *n*th position in the chromosome is assigned to the *k*th Mobile Harbor unit, where *k* is a remainder of *n*/*m*. When the remainder is zero, the job will be assigned to the *m*th Mobile Harbor unit. Figure 6 illustrates this procedure.

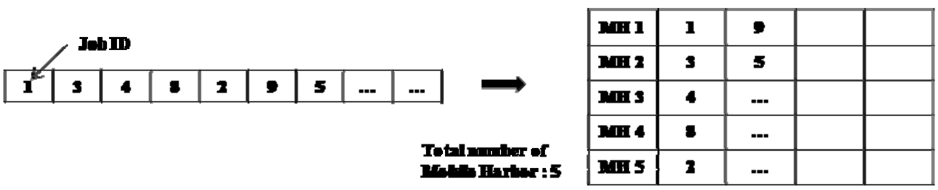


Figure 6. Means of generating a job sequence for each Mobile Harbor unit from a chromosome

4.2 Initial population

Before generating the initial population, we create null jobs which have zero processing time. When we generate a chromosome with only real jobs, every Mobile Harbor has the same number or a similar number of jobs compared to other Mobile Harbor units due to the method used to assign a job to a Mobile Harbor unit. This restricts the solution space of the GA algorithm. To explore a larger solution space, we create null jobs and generate a population using both real jobs and null jobs. By adding null jobs, it is possible to assign a fewer number of jobs to some of the Mobile Harbor units while others take a greater number of jobs.

In generating initial population, all jobs are sorted by a non-decreasing order of its ready time. Jobs from the same container ship have the same ready time. These jobs are randomly sorted. In addition, null jobs are assigned in the chromosome only after all real jobs are assigned. Consequently, positions of null jobs are always behind real jobs for each Mobile Harbor unit. Through this process, an initial population is generated by repeatedly creating chromosomes until it reaches the desired population size.

4.3 Objective value evaluation

As with the MIP model, the objective value for the GA problem is the sum of the staying times of each container ship. The staying time of a container ship is determined by the completion time of the last job of the container ship. For a given sequence of jobs, the completion time of each job is calculated. The completion time of job *i* depends on four factors: the ready time of job *i*, the conflict-delay time for job *i*, the semi- conflict-delay time for job *i*, and the availability of the Mobile Harbor unit job *i* is assigned to.

The ready time of job *i* is the arrival time of the container ship that job *i* belongs to. When job *i* is ready, say at time *t*, and the Mobile Harbor unit that job *i* is assigned to is available at the moment, job *i* can begin at time *t*. However, if there exists a job *j* at time *t* that is under the conflict condition with respect to job *i*, job *i* cannot begin until job *j* is completed and must wait for some time period. This time is referred to as the conflict-delay time. Another type of delay may arise due to the semi- conflict condition. When two jobs, job *k* and *l*, are under progress at time *t*, as shown in Figure 7, and job *i* is located in between job *k* and *l*, job *i* cannot begin because job *i* is in the semi- conflict condition with respect to both job *k* and *l* such that avoiding conflict with one job causes conflict with the other. In this case, job *i* must wait until either job *k* or *l* is completed. This delay is referred to as the semi- conflict-delay. The length of this delay is the minimum of the completion times of job *k* and *l*.

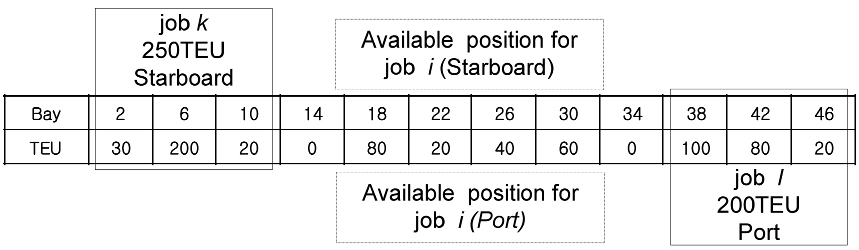


Figure 7. Conflicts between jobs in the semi- conflict condition

If job *i* is assigned to Mobile Harbor *k*, it can begin after completing the immediately preceding job, job *h*, of Mobile Harbor *k*. If Mobile Harbor *k* must visit a land-side berth to unload its containers, then Mobile Harbor *k*'s available time for job *i* is the completion time of job *h* plus the downtime of Mobile Harbor *k*. The downtime of Mobile Harbor *k* is the sum of the roundtrip travel time between the container ship and a land-side berth and the unloading time for Mobile Harbor *k*. If Mobile Harbor *k* has sufficient capacity to carry job *i* after the completion of job *h*, and thus do not need to visit the land-side berth, the downtime in this case is zero.

Based on these four factors, the completion time of job *i* is determined by equation (28),

$$\text{Completion time of job } i = max\{r_i, cf_i, scf_i, MH_i^k\} + \text{unloading time of job } i \quad (28)$$

, where r_i = the ready time of job *i*, cf_i = conflict-delay time of job *i*, scf_i = semi- conflict-delay time of job *i*, and MH_i^k = available time of Mobile Harbor *k* to operate job *i*.

4.4 Fitness function and Selection

The fitness value for a chromosome *i* is calculated based on a fitness function used in a quality-proportional roulette wheel selection [Moon (2008)], given by equation (29).

$$f_i = (C_w - C_i) + (C_w - C_b)/2. \quad (29)$$

C_w and C_b are the worst and best objective value, respectively, in the population at the current generation.

New population for the next generation is created as follows. First, the top 30% of the population in terms of the objective value is passed down to the next generation. Then the remaining 70% of the population is generated in two steps. At first, a preliminary population is generated by the roulette wheel selection where the selection probability is determined by the fitness value given by equation (29). Then, an order crossover (OX) is carried out [Jin (2000)]. A typical order crossover process is illustrated in Figure 8.

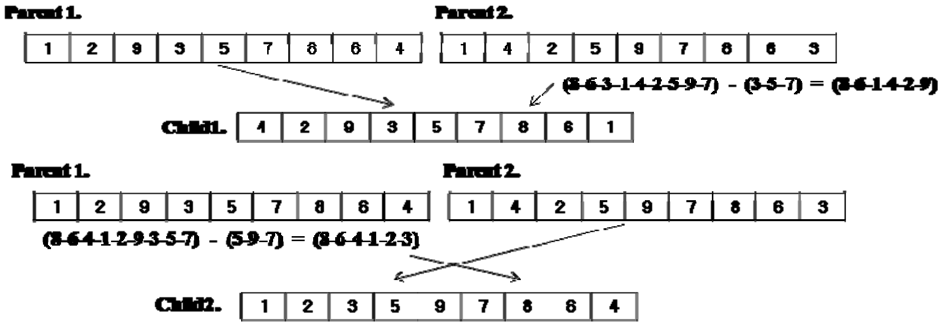


Figure 8. Order Crossover

5. COMPUTATIONAL RESULTS

A performance comparison between the MIP model and the proposed GA method is presented in this section. Optimal solutions by the MIP model are available only for a small-sized problem, and thus we conduct comparison between the solutions from the MIP model and the GA method for a relatively small-sized problem instances.

In this comparison, the GA method obtained an optimal value for the tested problem instances, that is the same as the MIP solution. Due to the small number of possible job sequences, it seems that the GA method is able to explore the solution space nearly exhaustively with the population size (up to 200) and the number of maximum generation (up to 1000) used in the experiments. The results are shown in Table 1. Computation time to acquire an optimal solution by the GA method was non-significant. We can conclude that the performance of the GA method is comparable to that of the MIP model for a small problem in terms of the quality of the solution with a shorter computation time.

Table 1. MIP model vs. GA method

Total number of container ships: 1						Total number of container ships: 2					
Exp. No	JOB M	H	MIP	GA	GAP[a]	Exp. No	JOB	MH	MIP G	A GA	P
1	6 2		1490	1490	0% 5		6	2	250 2	50 0%	
2	7 2		1490	1490	0% 6		7	2	2360 2	360 0	%
3	7 4		1550	1550	0% 7		7	2	1620 1	620 0	%
4	11 2		N/A	2942	- 8		10	4	N/A 1	290 -	

GAP[a] = 100 (%) · (final solution of GA - optimal solution of MIP) / optimal solution of MIP

For a large problem for which the MIP model practically cannot be solved, the GA method does not guarantee the optimality of a solution it produces. To improve the quality of a near-optimal solution, we may choose to increase the number of generations. Yet, this does not necessarily improve the quality of the obtained solution due to the premature convergence of the population. To address this problem, we conduct experiments to identify appropriate values for the population size and the crossover probability since they influence the quality of final solutions. Due to the probabilistic nature of a Genetic Algorithm, each run generates different solutions, and thus we run a multiple experiments for a given set of population size and crossover probability.

The experiments are designed to test two factors, population size and crossover probability, at three levels. A test problem has 4 container ships, 30 jobs, and 5 Mobile Harbor units. Each case of the experiments is replicated five times, each with the maximum generation number of 1,000. The results of the experiments are presented in Figure 9. As shown in Figure 9, it seems that crossover probability of 0.5 produces the best result and that a larger population size is desirable than a smaller population size for the Mobile Harbor scheduling problem.

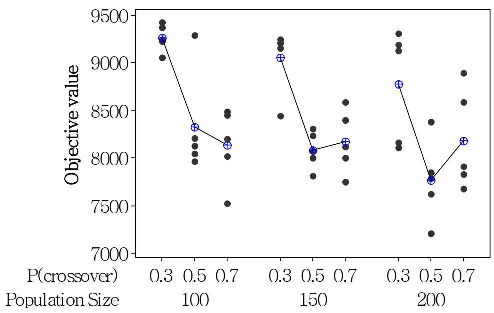


Figure 9. The effect of Population size and Crossover Probability

6. DISCUSSION

While the MIP model guarantees an optimal solution for the Mobile Harbor scheduling problem, it can be solved practically only for a small sized problem. In our case, we were not able to obtain solutions by the MIP model even for a relatively small number of jobs, e.g. 10. In contrast to the MIP model, the GA method consumes little computation time to obtain a near-optimal solution. It even produces an optimal solution for small sized problem instances. One of the reasons why the GA method can find high-quality solutions much quicker than the MIP model is the way an idle time of a Mobile Harbor unit is handled. An idle time is related to a time window within which a Mobile Harbor can start its assigned job any time without affecting the final job completion time. In the GA method, when there is such a case, the Mobile Harbor is forced to start its assigned job at the earliest possible time. As such, for a given job sequence, an objective function – total ships' staying time – can be evaluated quickly. On the other hand, in the MIP model, it evaluates all the possible job starting time in the available time window, which causes much longer evaluation time.

Quality of a final solution obtained from the GA method is largely influenced by its crossover probability and population size. As shown in Figure 9, it is suggested that 0.5 be the most appropriate crossover probability. Crossover probability of 0.3 seems to generate insufficient number of crossover events. With 0.3, the method may not explore large enough solution space, and converge to a solution prematurely. With a higher probability of crossover at 0.7, the GA method does not seem to produce a better solution, even though it probably explores a larger solution space. This may be because a solution with a chance to becoming the best solution does not stay long in a population due to a strong degree of diversification with a high crossover probability. In fact, even after 1,000 generation, its population is far from being converged. It is thus conjectured that increasing the number of generation is likely further increase the quality of the solution.

A final solution's quality is improved as the population size is increased, which is intuitively expected. While this positive effect of population size is intuitive, an interesting finding is that a final solution appears at an earlier generation for a larger population than a smaller population. In other words, when we calculated a computation time (i.e. generation number) taken from the start to the generation when a final solution first appeared, a run with a larger population usually obtains a final solution earlier than a run with a smaller population. These findings suggest that, for a given computational load, a larger population with a smaller number of generations is a more efficient scheme than a smaller population with a larger number of generations.

In addition to identifying appropriate parameter values, further study to develop an effective initial solution is also needed. A final solution of the GA method is influenced by an initial solution. As such, initial solutions generated by priority rules such as the LPT (Longest Processing Time) or hybrid priority rules can enhance the quality of the final solution of the method as well as the convergence time to the final solution [Lee *et al*. (2010)].

7. CONCLUDING REMARKS

A Mobile Harbor system is a new concept for transporting containers between a container ship and a land-based port, and it introduces scheduling issues unique to its operational characteristics. This paper studies a Mobile Harbor scheduling problem as an m-parallel machine with limited capacity where jobs have conflicts with each other. For the Mobile Harbor scheduling problem, this paper presents a MIP formulation that is a modified form of a quay crane scheduling problem. Solving the MIP model requires considerable computation time, and it is solvable, in a practical sense, only for small sized problem instances. To overcome this computational limit, a heuristic method using GA is proposed as an alternative. The proposed GA method appears to perform very well for a small sized problem. A desirable set of parameter values for the GA method are also proposed, albeit derived from a limited set of test cases.

The methods studied in this paper can be applied to a system whose operational characteristics show similarity with a Mobile Harbor system. For example, a quay crane scheduling problem at a slip berth [Chae *et al*. (2007)] can be solved by applying the model developed in this study.

ACKNOWLEDGEMENTS
This study was supported by the Korean Ministry of Knowledge Economy.

8. REFERENCES

Chae, JW., Park, WS., Yi, JH., Jeong, G. (2007). A floating mobile quay for a very large container ship. *Asian and Pacific Coasts.*

Jin, GG. (2000). *Genetic Algorithms and Their Applications*. KyoWooSa, Korea.

Kim, KH., Park, YM. (2004). A crane scheduling method for port container terminals. *European Journal of Operational Research, 156:* 752-768.

Lee, D.H., Wang, H. Q., Miao, L. (2008). Quay crane scheduling with non-interference constraints in port container terminals. *Transportation Research Part E: Logistics and Transportation Review, 44:* 124-135.

Lee, T., Sung, I., Nam, H. (2010). An Extended M-parallel Machine Problem with a Limited Machine Capacity and Interference between Jobs. *Proceeding of the Spring joint Conference of Korean Institute of Industrial Engineers and Korean Operations Research and Management Science Society.*

Moccia, L., Cordeau, J. F., Gaudioso, M., Laporte, G. (2006). A branch-and-cut algorithm for the quay crane scheduling problem in a container terminal. *Naval Research Logistics, 53:* 45-59.

Moon, BR. (2008). *Genetic Algorithms for beginners*. Hanbit media, Korea.

Sammarra, M., Cordeau, J. F., Laporte, G., Monaco, M. F. (2007). A tabu search heuristic for the quay crane scheduling problem. *Journal of Scheduling, 10:* 327-336.

Session H3: Seaport and Transportation 8

·Day2: Sep. 16 (Thu.)

·Time: 14:40 - 16:00

·Chair: Herbert Kopfer

·Room: Iris, 4F

CONTAINER REHANDLING AT MARITIME CONTAINER TERMINALS

Marco Caserta[1], Silvia Schwarze[1], and Stefan Voß[1*]

[1] University of Hamburg
Germany
e-mail: caserta@econ.uni-hamburg.de
schwarze@econ.uni-hamburg.de
stefan.voss@uni-hamburg.de

Abstract: In this paper, we review recent contributions dealing with the rehandling of containers at maritime container terminals. The problems studied in the paper refer to a post-stacking situation, i.e., problems arising after the stacking area has already been arranged. In order to increase efficiency of loading/unloading operations, once updated information about the state of the containers as well as of the vessels become available, it is possible to reshuffle the container yard, or a portion of it, in such a way that future loading operations are carried out with maximal efficiency. The increase in efficiency of loading/unloading operations has a bearing on the berthing time of the vessels, which, in turn, is a widely accepted indicator of port efficiency. Three types of post-stacking problems have been identified, namely (i) the remarshalling problem, (ii) the premarshalling problem, and (iii) the relocation problem. With respect to each one of these problems, a thorough explanation of the problem itself, its relevance and its connections with other containers handling issues are offered. In addition, algorithmic approaches to tackle such problems are summarized in the paper and some new approaches are proposed.

ALIGNING BARGE AND TERMINAL OPERATIONS IN THE PORT OF ROTTERDAM BY MEANS OF A MULTI-AGENT SYSTEM

Albert Douma[1] and Henk Zijm[1]

[1]Department of Operational Methods for Production and Logistics
School of Management and Governance
University of Twente
Enschede, The Netherlands
E-mail: a.m.douma@utwente.nl

Abstract: In this paper we present policies that are shown to significantly improve the alignment of barge and terminal operations in the port of Rotterdam. This alignment is important, since barges (transporting containers between the port and the hinterland) visit on average seven terminals in the port to (un)load containers. The inefficiencies, resulting from a poor alignment, lead to significant (in)direct costs and affect the competitiveness of the Port of Rotterdam. Central coordination is not acceptable for companies for competitive reasons. In an earlier study we developed a distributed planning approach. We apply this approach to a realistic model of the Port of Rotterdam in the period 2006-2007. Our aim is to indicate the impact of a distributed planning approach, also in comparison to central coordination. We do not aim to quantify the exact impact of our approach. The results indicate that distributed planning is very promising in the Port of Rotterdam and performs well compared to central coordination.

1. INTRODUCTION

Barge container transport is of vital importance for ports, like Antwerp and Rotterdam, to transport containers to the hinterland and back. Barge container transport is attractive compared to other modalities, since it is less harmful for the environment, can relieve the congested rail and road infrastructure, and has lower transport tariffs. Nowadays, port authorities even force companies to increase the share of containers transported by barge. However, over the last decade, barges suffer more and more from long waiting times at terminals in the port. Important causes are a poor alignment of barge and terminal operations and the fact that appointments about handling times are uncertain. The result is an underutilization of terminal and barge capacity. The problem is urgent in Rotterdam and Antwerp, and will become more important in view of increasing container flows. We call this problem the *barge handling problem*.

The barge handling problem concerns the optimization of the alignment of barge and terminal operations in the port. This is a complicated matter for several reasons. First, in the Port of Rotterdam, barges usually visit about three to fifteen terminals to load and unload containers (see Figure 1 for a picture of the Port of Rotterdam). Disruptions propagate quickly through the port causing dominoes effects. Second, barge operators compete with each other (and so do terminal operators). This results in two important constraints regarding a solution; i) companies want to stay autonomous, i.e., in control of their own operations, and ii) companies are reluctant to share information that could possibly undermine their competitive position. Third, the barge handling problem is a dynamic problem, meaning that a lot of events and (small) disruptions occur which may require reconsideration of terminal and barge plans.

In this paper we mainly talk about barge and terminal operators. Barge operators are companies that offer and organize barge container transportation services to and from the hinterland. These companies usually do not own barges themselves, but contract barge companies. Terminal operators are companies that operate a terminal.

1

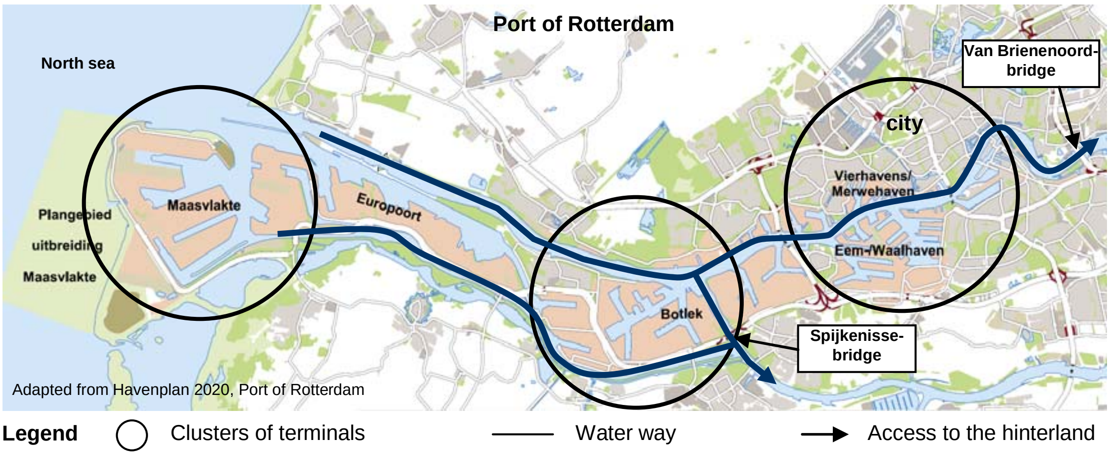


Figure 1. Map of the Port of Rotterdam. Highlighted are the important container terminal clusters, the main waterways, and the hinterland access points

Nowadays, barge and terminal operators try to align their activities by making appointments. These appointments are made by telephone, fax, and e-mail. Unfortunately, it happens frequently that appointments are not (or cannot be) met by either the barge or the terminal operator. The result is that barge operators face uncertain waiting and handling times at terminals, and terminal operators deal with uncertain arrival times of barges.

In the past attempts have been made to coordinate terminal and barge activities centrally. However, this solution turned out to be not acceptable. Companies would loose their autonomy and have to share a lot of information. Therefore, in the past another approach has been proposed (Connekt 2003; Melis et al. 2003; Schut et al. 2004), namely distributed planning by means of a Multi-Agent system. In an earlier study we developed and evaluated a Multi-Agent system for the barge handling problem (Douma et al. 2009). This Multi-Agent system has been tested on fictitious port settings to investigate how certain agent' interaction protocols perform. Additionally, we compared the performance of our Multi-Agent system with central coordination.

To the best of our knowledge, the barge handling problem has, except for the studies mentioned, not been studied in the literature before. The problem is, however, related to problems in different fields. We mention the berth allocation problem (Cordeau et al. 2005; Stahlbock and Voss 2008), the ship routing and scheduling problem (Christiansen et al. 2004), the attended home delivery problem (Campbell and Savelsbergh 2006), the hospital patient scheduling problem (Decker and Li 2000), and multi agent theory (Wooldridge and Jennings 1995). In Douma et al. (2009) we provide a discussion of each of these fields. For the role operations research methods play in the in the optimization of terminals operations we refer to an extensive literature study by Steenken et al. (2004) and Stahlbock et al. (2008).

The contribution of this paper is to investigate how the Multi-Agent system developed at the University of Twente performs in a realistic setting of the Port of Rotterdam. We therefore use the models developed in previous papers and we apply these to a realistic model of the Port of Rotterdam in the period 2006-2007. The outline of the paper is as follows. In Section 2 we discuss the notion of distributed planning and Multi-Agent systems. Section 3 describes in more detail the interaction protocol used by the agents, namely service-time profiles. In Section 4 we describe the intelligence of the agents and in Section 5 we discuss our approach to evaluate the performance of our Multi-Agent system through relevant performance indicators. Section 6 describes the simulation model and the off-line benchmark, resembling central coordination. Section 7 describes the performance of our Multi-Agent system in the Port of Rotterdam. Finally we draw our conclusions in Section 8.

2. DISTRIBUTED PLANNING/MULTI-AGENT SYSTEMS

Distributed planning (or decentralized control) is the coordination of activities among multiple self-interested parties that make decisions independently and where the system performance is a result of the actions and decisions of individuals. Centralized control on the other hand is the coordination of activities by a single party that decides on and coordinates the actions of individuals, to maximize the system performance.

Distributed planning can be realized by means of a Multi-Agent system. A Multi-Agent system is a system consisting of multiple intelligent (software) agents. An intelligent agent is usually defined as a software based object with key properties: autonomy, social ability, reactivity, and pro-activeness (Wooldridge et al. 1995). The purpose of an agent is to take over tasks of its principal, i.e., the human or organization it represents.

A Multi-Agent system is a promising approach to the barge handling problem, since it can do justice to the specific business constraints. Every party is equipped with a software agent which acts on its (the principal's) behalf. In a Multi-Agent system there are two important elements, namely i) the interaction protocol, and ii) the strategy of players.

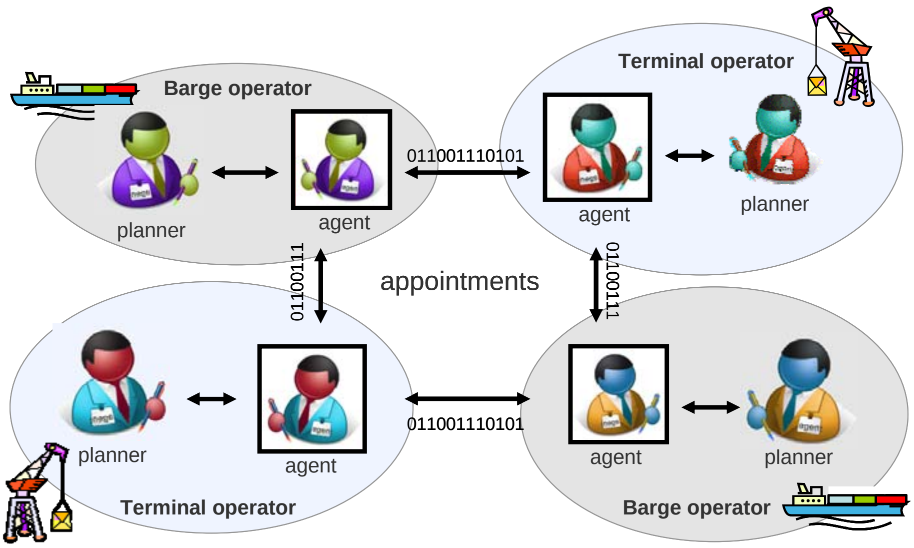


Figure 2. A visualization of a Multi-Agent system with two barge and two terminal operator agents

In our Multi-Agent system we distinguish to types of agents, namely barge and terminal operator agents (see also Figure 2). In the Multi-Agent system we imitate more or less the way the market is currently structured to make the solution acceptable for the players involved. For the same reason we assume that the strategy of an agent is to behave opportunistically, i.e., exploiting opportunities for its principal with no regard for the consequences of others. We describe the interaction protocol we apply in the following section.

3. SERVICE-TIME PROFILES

Starting point in the design of our Multi-Agent system is i) the aim to create a real-time planning system and ii) the assumption that barges and terminals want to make (reliable) *appointments* for the transshipment of containers. Making appointments has several advantages, since it gives barge and terminal operators certainty in their operations.

We propose the following definition of an appointment. An appointment made by a barge and a terminal is an agreement from two sides. The barge promises the terminal to be present at the terminal at a latest arrival time. The terminal in turn guarantees the barge a latest departure time, if the barge keeps its promise. By making appointments, the barge uses the guaranteed latest departure times at preceding terminals to calculate the arrival time at the next terminal.

We divide the communication between the barge operator and terminal operators in two phases (see Figure 3). In the first phase, the barge operator agent attempts to obtain information about the occupation of each terminal to be visited. Based on this information it determines its preferred rotation (sequence of terminal visits). In the second phase, the barge operator agent makes appointments with the terminal operator agents, based on the rotation determined in the first phase.

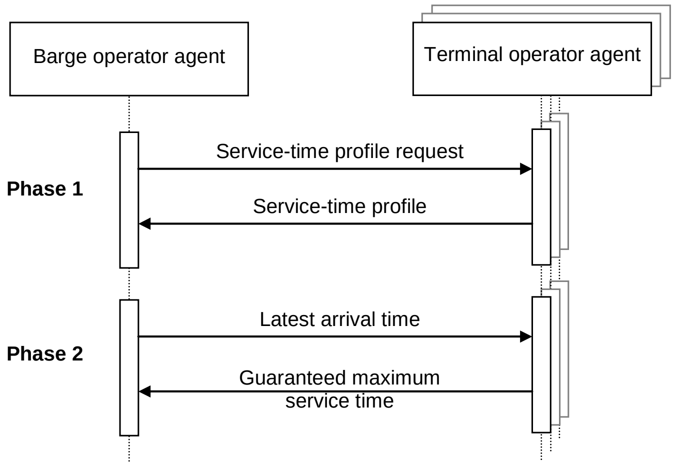


Figure 3. A sequence diagram of the communication between a single barge operator and multiple terminal operators in our model

In the first phase terminals hence reveal information on their occupation. They do so by issuing, so-called, service-time profiles. A service-time profile is barge and time specific and denotes the maximum service time (sum of waiting time and handling time at a terminal) for a barge for every possible arrival time during a certain time period. The maximum service times determining the service-time profile are guaranteed maximum service times. In Figure 4 we give an example of a service-time profile for a specific terminal. The terminal is closed from time 22:00 until 6:00. During this time no transshipments activities take place. If the handling of a barge is not completed before closing of the terminal, then the handling is preempted and resumed again on opening of the terminal. In Figure 4 one can see the maximum service time given a certain arrival time in a certain time period. For instance, when the barge promises to arrive not later than time 14:00 then its maximum service time is equal to 2 hours.

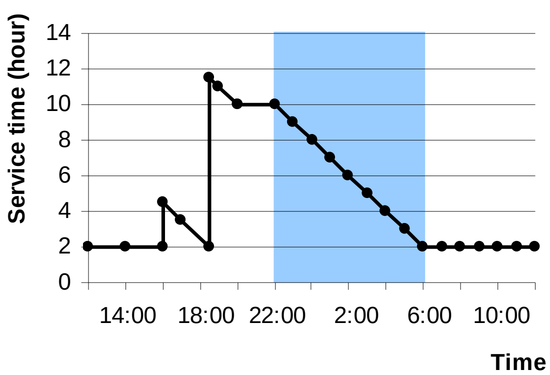


Figure 4. An example of a service-time profile. Closing of the terminal (from 22h-6h) is denoted by the shaded area

In addition to the service-time profiles we also consider other levels of information exchange in the first phase of the barge-terminal communication. We define a level of information exchange as the extent to which a terminal gives insight to a barge in the occupation of the terminal during the day. In our study we consider (apart from service-time profiles) the following two levels of information exchange:

1. *No information*. Terminals reveal no information about their occupation. The barge operator therefore determines a rotation by finding the shortest path along all the terminals concerned.
2. *Yes/No*. A barge operator can ask terminals repeatedly whether a certain arrival time is convenient and the terminal replies only yes or no. To find 'the best' rotation, the barge operator may delay terminal visits or select an alternative rotation. The basic idea is that barges retrieve information about the occupation of terminals by asking whether a time slot is available. This means that the information revealed by the terminals gives only limited insight in its occupation. This level of information exchange resembles to a certain extent the interaction protocol implemented in a preceding project, called, APPROACH 1 (Connekt 2003).

We stress that the levels of information exchange, varying from "No information" via "Yes/No" to "Service Time Profiles", relate to the first phase of the interaction protocol. The second phase of the interaction protocol, where barge and terminal operator agents make appointments, is straightforward, regardless the level of information exchange in the first phase.

4. BARGE AND TERMINAL OPERATOR INTELLIGENCE

Let us briefly describe the barge and terminal operator intelligence in our Multi-Agent model. The aim of the barge operator is to minimize the sojourn time of a barge in the port. To do so, the barge solves a time-dependent traveling salesman problem (TDTSP) using the following information: i) the set of terminals that has to be visited, ii) the sailing time between each pair of terminals, and iii) (time dependent) service-times at terminals derived from the service-time profiles. For solving the TDTSP we use depth-first branch-and-bound for rotations with at most seven terminals and a DP-heuristic (developed by Malandraki and Dial 1996) for rotations with more than seven terminals (see Douma 2008).

The terminal operator in turn has to decide when a barge is or can be handled. We assume that terminal operators treat all barges equal, which means that a terminal accepts an appointment with a specific barge if the barge can be handled. Terminal operators (re)schedule their quays using list-scheduling in combination with dept-first branch-and-bound. For the scheduling of the barges we apply list-scheduling in combination with depth-first branch-and-bound (see Douma 2008).

5. PERFORMANCE EVALUATION

The performance of the Multi-Agent system is evaluated by means of simulation. In the simulation we compare the performance of the different levels of information exchange. Additionally, we use a centralized approach to see how well our Multi-Agent system performs compared to central coordination. This centralized approach is an off-line benchmark that solves the problem centrally (for a single objective function: minimize the total tardiness) and assumes all information over the whole planning horizon to be known in advance. To facilitate a sound comparison of the two models we use the same data sets. The database contains a complete set of all the relevant parameters describing a scenario. Note that a difference in performance between our Multi-Agent model and the off-line benchmark will be partly due to the fact that in the off-line benchmark information over the planning horizon is known in advance.

We use barge related performance indicators to evaluate the performance of both approaches. The reason is that the utilization of a terminal is determined by its available capacity and the workload. Both cannot be influenced during the simulation. However, by processing these barges efficiently, the terminal can reduce the average waiting times at the terminal and thus influence the average sojourn time of barges in the port. The indicators we use are derived from scheduling literature and are related to the off-line benchmark. The key performance indicators we use are:

1. *Average lateness.* This indicator gives an impression of how 'early' or 'late' barges leave the port on average. Lateness is the actual time the barge leaves the port minus the time set in the sailing schedule. Lateness can be negative.

2. *Average tardiness*. The tardiness of a barge is equal to the maximum of zero and the lateness of the barge. The lateness is equal to the actual time the barge leaves the port minus the time set in the sailing schedule. Note that lateness can be negative.

The time available for a barge to complete its activities in the port is derived from its sailing schedule. A sailing schedule is sometimes determined a year in advance. In our study we consider two extreme variants, namely fixed and variable time windows. A fixed time window means that all barges have the same time window irrespective of the number of terminals they have to visit. A variable time window means that each barge has an individual time window depending on its rotation length. The fixed time window is calculated based on the expected total sailing and handling time needed by an average barge to complete all its activities in the port, multiplied with a slack factor to compensate for waiting times. The variable time window is calculated based on a *fixed* percentage of slack per rotation and a slack factor per terminal (and hence varies per rotation, dependent on the number of terminals to be visited).

6. SIMULATION MODEL AND THE OFF-LINE BENCHMARK

The simulation model we have made is a discrete event simulation. Event based means that after an event the barge or terminal can perform an action resulting in a state transition, after which the state of the system remains unchanged until the next event (Law and Kelton 2000). A terminal (or a terminal quay) can be in one of the following states: being idle, handling a sea vessel, handling a barge, or being closed. For a barge we consider the following states: sailing, waiting, handling, and mooring. Agents are modeled as decision making objects which can be invoked during the simulation.

For the off-line benchmark we use a Resource Constrained (Multi-)Project Scheduling Problem approach (RCPSP; see Demeulemeester and Herroelen 1992) as a base model. In the basic RCPSP, a single project has to be scheduled, which consists of a number of activities. These activities have to be processed on a number (not necessarily all) of the available resources, consisting of a number of parallel processors. In the RCPSP, terminals are represented as resources with a number of parallel resources equal to the number of quays. Barges are represented as independent projects. Each project consists of a set of activities (terminal visits). Opening times of terminals are denoted in so-called resource profiles, specifying the available capacity over time. The primary objective in the off-line benchmark is to minimize the total tardiness and the secondary objective to minimize the fraction of barges leaving the port late. For a more detailed description of the off-line benchmark see Douma et al. (2009).

7. PORT OF ROTTERDAM

In this section we study the performance of our Multi-Agent system in the Port of Rotterdam for the period 2006-2007. Historical data about the performance of terminals and barges in that period is hard to get. Data about the port operations are limited, ambiguous, incomplete, distributed, and sometimes even contradicting. Additionally, structural changes in the port take place frequently, which means that a model of the port today may not be valid anymore tomorrow. Nevertheless, we think that we managed to develop a realistic model of the Port of Rotterdam which can provide valuable insights in the performance of our Multi-Agent system. We obtained data from experts, the Port of Rotterdam, company information, and other sources such as interviews and master theses.

At the start of 2007, there were 33 terminals in the Port of Rotterdam that handle containers. These terminals are distributed in clusters over the port (see Figure 1). Each terminal has its own characteristics, i.e., the number of quay-crane-crew combinations, opening times, average call sizes, utilization degree, and whether they handle sea vessels. Typically, terminals in the city area are closed during night hours, whereas terminals on the Maasvlakte are opened 24 hours a day. Barges visit on average seven terminals. The number of arriving barges each day is derived from the capacity and utilization degrees of terminals. The number of arriving barges each day is about 50-60. Travel times between terminals are based on real distances. Note that travel times between clusters are relatively large (compared to travel times within clusters), which implies that most barges (entering the port via the “Van Brienenoord-bridge”) visit the clusters in the following order: City – Botlek – Maasvlakte – Botlek – City.

Based on the data about terminals and barges we developed a base case. The base case is a realistic representation of the activities in the Port of Rotterdam in the period 2006-2007. The base case does not allow us to quantify the impact of implementing a Multi-Agent system on, e.g., the operations of terminals and barges or the economical impact on the port. Apart from the reasons mentioned above about the quality of the data we also ignore, e.g., stochastic influences and specific preferences or working practices of barge and terminal operators. However, the base case gives us insight in the performance of our Multi-Agent system, and how it compares with central coordination.

We simulated the base case and compared the results with the off-line benchmark. Figure 6 compares the performance of the different levels of information exchange and the off-line benchmark in terms of average tardiness. The results indicate that our Multi-Agent system (with service-time profiles) performs well in comparison with the off-line benchmark. Additionally, we see that other levels of information exchange (no information and yes/no) perform significantly worse. If we look at the average lateness (see Figure 7) we find similar results.

The average sojourn time of a barge in the port (in case of Service-Time profiles) is 36 hours. This consists of about 9 hours sailing, 9 hours handling, and 18 hours waiting. The average waiting time boils down to 2.6 hours per terminal. The sojourn time of barges (in case of No-Information) is about 69 hours, consisting of 7 hours sailing, 9 hours handling, and 54 hours waiting. As one can see, the increased sojourn time in case of No-Information, is mainly due to a decrease in the average waiting time. The fact that No-Information on average has less sailing time is explained by the fact that the barge operator has no information about waiting or service times and just minimizes the total sailing time. In case of service-time profiles the barge operator weighs the sailing and the service time when determining a rotation.

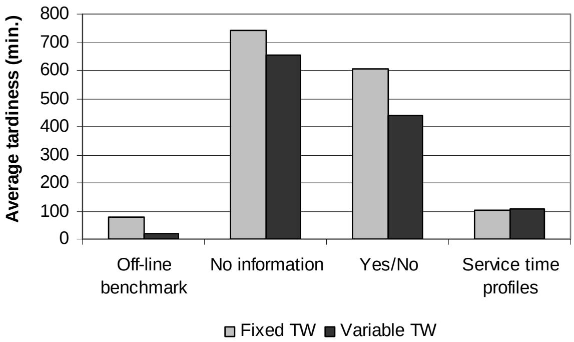


Figure 6. Average tardiness for different levels of information exchange in the Multi-Agent system compared with the off-line benchmark

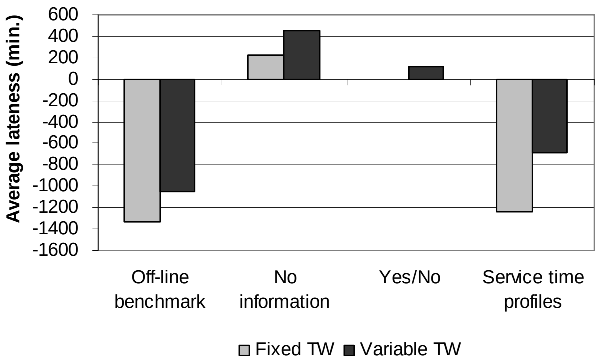


Figure 6. Average lateness for different levels of information exchange in the Multi-Agent system compared with the off-line benchmark

If we analyze the rotation planning of barge operators in more detail, we find that there is a clear relation between the order in which terminals are visited and the waiting time at these terminals at the time the barge plans its rotation. The results show that bottleneck terminals are visited later in the rotation. In the waiting time of the bottleneck terminals (the terminals with the longest waiting times) other terminals are visited. This means that the sojourn time of a barge in the port not necessarily decreases when it visits less terminals.

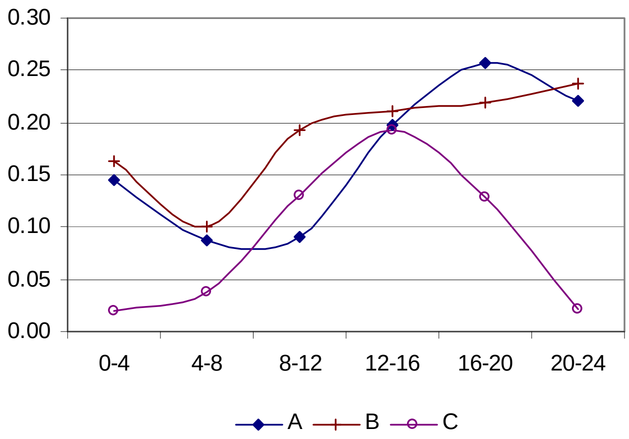


Figure 7. Fraction of barges arriving at terminals A, B, and C during certain time intervals of the day

Another result regards the arrival pattern of barges at terminals. In Figure 8 we depict the arrival pattern of barges during the day for terminals with different opening times and at different port locations. If we look at terminal C (closed from 10 p.m. to 7.30 a.m.) we can see that this terminal has clearly more barge arrivals during the day than in the night. Terminal B, however, is opened 24h a day and we see that especially at the start of the day there is a decline in the fraction of barge arrivals. The reason is that this terminal is surrounded by terminals that close during the night. When these terminals open, barges prefer to head to these terminals first. Terminal A, on the other hand, is opened 24h and not surrounded by terminals that close during the night. One can see that the arrival pattern has a wave-like shape during the day. The reason for this is that at the end of the day, several terminals in other regions close and that barges prefer to visit 24h terminals during the night.

8. CONCLUSIONS

In this paper we discussed the barge handling problem in the Port of Rotterdam. The barge handling problem concerns the alignment of container barge and terminal operations in a port. Barges are used to transport containers from the port to the hinterland, and vice versa. Every time a barge arrives in the port, it visits several terminals to load and unload containers. The sequence in which a barge visits these terminals depends, among others, on the availability of terminals. The availability of terminals, in turn, is depending on the other barges and the sea vessels that have to be processed.

The barge handling problem is not easy to solve due to several complicating factors, like the fact that parties want to stay autonomous and are reluctant to share information. Previous attempts to provide a solution to the barge handling problem made clear that a centralized solution approach is not acceptable for the actors concerned, and that a decentralized planning approach may be one of the few ways to solve the problem. The problem is highly relevant to many sea ports such as Rotterdam. The inefficiencies, resulting from a poor alignment of barge and terminal operations, lead to significant (in)direct costs and also affect the attractiveness of the Port of Rotterdam as a node in global container transportation chains.

In this paper we described the Multi-Agent system we developed for the barge handling problem. We discussed the interaction protocol used (based on service-time profiles), the intelligence of the barge and terminal operator agent, and the way we evaluate the performance of our Multi-Agent system. Additionally we provided results on the performance of our Multi-Agent system in the Port of Rotterdam. These results indicate that a Multi-Agent system with service-time profiles performs well compared to the off-line benchmark, and outperforms the other 'levels of information exchange'. The comparison with the off-line benchmark suggests that our Multi-Agent system is promising in the Port of Rotterdam.

ACKNOWLEDGEMENT

This research is performed within a TRANSUMO program. TRANSUMO is a Dutch foundation which supports research to realize a transition to sustainable mobility. We are grateful to TRANSUMO for financing this research.

9. REFERENCES

Campbell, A. M., and Savelsbergh, M. (2006). Incentive schemes for attended home deliveries. *Transportation Science, 40*(3), 327-341.

Christiansen, M., Fagerholt, F., and Ronen, D. (2004). Ship Routing and Scheduling: Status and Perspectives. *Transportation Science, 38*(1), 1-18.

Connekt. (2003). *Eindrapport Approach*. Delft. Report is in Dutch: Connekt.

Cordeau, J. F., Laporte, G., Legato, P., and Moccia, L. (2005). Models and tabu search heuristics for the berh-allocation problem. *Transportation Science, 39*(4), 526-538.

Decker, K., and Li, J. (2000). Coordinating Mutually Exclusive Resources using GPGP. *Autonomous Agents and Multi-Agent Systems, 3*(2), 133-157.

Demeulemeester, E. L., and Herroelen, W. S. (1992). A branch-and-bound procedure for the multiple resource-constrained project scheduling problem. *Management science, 38*(12), 1803-1818.

Douma, A. M. (2008). *Aligning the Operations of Barges and Terminals through Distributed Planning.* PhD thesis, University of Twente.

Douma, A. M., Schutten, J. M. J., and Schuur, P. C. (2009). Waiting profiles: an efficient protocol for enabling distributed planning of barge rotations along terminals in the Port of Rotterdam. *Transportation Research Part C, 17*(2), 133-148.

Law, A. M., and Kelton, W. D. (2000). *Simulation Modeling and Analysis* (Vol. 3). Singapore: McGraw-Hill.

Malandraki, C., and Dial, R. (1996). A restricted dynamic programming heuristic algorithm for the time dependent traveling salesman problem. *European journal of Operational Research, 90*, 45-55.

Melis, M., Miller, I., Kentrop, M., van Eck, B., Leenaarts, M., Schut, M., et al. (2003). *Intelligent Logistics Concepts.*

Schut, M., Kentrop, M., Leenaarts, M., Melis, M., and Miller, I. (2004). APPROACH: Decentralised Rotation Planning for Container. *Proceedings of the 16th European Conference on Artificial Intelligence.*

Stahlbock, R., and Voss, S. (2008). Operations Research at Container Terminals: a literature update. *OR Spectrum, 30*, 1-52.

Steenken, D., Voß, S., and Stahlbock, R. (2004). Container terminal operation and operators research - a classification and literature review. *OR Spectrum, 26*, 3-49.

Wooldridge, M. J., and Jennings, N. R. (1995). Intelligent Agents: Theory and Practice. *The Knowledge Engineering Review, 10*(2), 115-152.

COLLISION-SAVE ROUTING OF A FLEET OF AUTOMATED GUIDED VEHICLES THROUGH A NETWORK: THE FORWARD-POSITIONING PLANNING PROBLEM

Herbert Kopfer[1], Jörn Schönberger[2], Kap Hwan Kim[3] and Su Min Jeon[4]

[1]Chair of Logistics
University of Bremen
Wilhelm-Herbst-Straße 5
28359 Bremen, Germany
E-Mail: kopfer@uni-bremen.de

[2]Operations Research and Supply Chain Management
RWTH Aachen
Templergraben 64
52059 Aachen, Germany
E-Mail: schoenberger@or.rwth-aachen.de

[3]Department of Industrial Engineering,
Pusan National University
Jangjeon-dong, Kumjeong-ku
Busan 609-735, Korea
E-mail: kapkim @pusan.ac.kr

[4]Faculty of Mechanical, Maritime and Material Engineering
Delft University of Technology
Mekelweg 2, 2628 CD
Delft, The Netherlands
E-mail: S.Jeon @tudelft.nl

1. INTRODUCTION

Container Terminals are critical components of today's world-wide supply networks. The shipping and transportation activities preceding and following the intra-terminal handling exploit significant economies of scale by moving large numbers of containers simultaneously (especially by ship). However, each single container requires an individual handling in a terminal because each container has an individual destination or source and each single container has to be treated individually. The exploitation of economies of scale is limited within a container terminal. It is tried to increase the reliability of terminal operations as well as to reduce the costs for the operations by lifting the degree of automation of terminal operations.

The transportation of empty or filled containers from their specific terminal entry points to their terminal exit points is the most time and workforce consuming logistics activity in a container terminal. Automated Guided Vehicles (AGV) are deployed to perform these operations without the need of having a human driver on a vehicle. One of the critical issues that decides about the benefit and success of an AGV deployment is an effective, efficient and reliable control strategy for coordinating and planning the AGV movements in a terminal (Kim and Bae, 2004). A major aspect in the AGV routing is process security. Although the AGVs operate in a man-free area it is necessary to take special efforts in the planning of AGV operations in order to prevent accidents among the AGVs. Therefore, buffers between moving AGVs have to be considered for the planning of AGV-routes in a terminal (Möhring et al., 2005).

In this contribution, we propose a specific approach to generate collision-preventing routes of vehicles in a terminal network. This planning approach is based on a sequentialization of the AGVs to be routed. Whenever a vehicle is routed, it blocks some transfer points in a network for other AGVs for a certain time in order to keep a least distance among moving AGVs.

2. ROUTING AGVS THROUGH A TERMINAL

The operational control of AGV-movements in a (terminal) network comprises the following three major decision tasks (Möhring et al., 2005). (i) Task-assignment addresses the assignment of transportation tasks to the available AGV. (ii) Subject of AGV-routing is the determination of a path along which an AGV travels through the given network in order to fulfill a previously assigned task. (iii) The avoidance of collisions between AGVs is the third aspect of the AGV control.

A network connects the major facilities within a terminal or production system by providing transportation infrastructure. In a terminal, commodities (e.g. containers) are moved, stored and transshipped. A typical example is a container terminal in a seaport where the commodities are the filled and empty containers that are brought or shipped away by short sea feeder ships, deep-sea ships or barges.

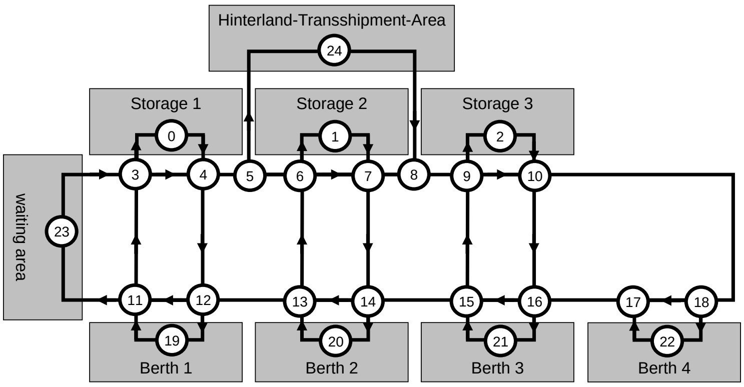


Figure 1. Example of an AGV-operated terminal

An example terminal layout is presented in Figure. 1. We have four kinds of facilities (gray shaded items). In the waiting area, idle AGVs wait for being assigned to a task. In a berth area transshipments from an AGV to a ship (or vice versa) are carried out. In a Hinterland-Transshipment-Area commodities that are coming into the terminal by truck or railway or that leave the terminal by truck or train are transshipped from a AGV to another means of transport and vice versa. Commodities that have entered the terminal but which do not leave it immediately are stored for a certain period in a storage area. Here transshipments from an AGV to a storage lot are executed.

A limited number of AGVs is moving between the different areas along lanes. These lanes are exclusively used by AGVs in order to prevent accidents. They are represented by the black lines in Figure. 1. An AGV travels along lanes at constant speed. At some designated positions (nodes) in the lane network the AGVs can perform additional activities. In a waiting area (node 23) an AGV waits if it is not executing some tasks. If an AGV is in a berth area (nodes 19-22) the AGV is loaded or unloaded by a special crane. It is loaded with some commodities (i.e. a container) just lightered from a ship or (if the AGV has brought a container to the berth area) the container is unloaded from the AGV and loaded into a ship. Only one AGV can be situated in such a node at a certain time point. Similar properties hold for the storage areas (nodes 0-2) and the Hinterland-Transshipment-Area (node 24). At the crossing of two lanes the AGV can change its direction and change the used lane (among other nodes: 3, 4, 5, 6, 7 …).

In a terminal consisting of the aforementioned nodes and the connecting arcs the AGVs are used to fulfill tasks by traveling along paths in this network. A task represents the instruction to travel (loaded or empty) from an origin node to a given destination node using the lane-network. We can distinguish two types of tasks (i) transportation tasks and (ii) repositioning tasks. If an AGV executes a transportation task (n_1,n_2) then it is instructed to travel from n_1 to n_2. At node n_1 it has to pickup a commodity and it has to delivery this commodity to n_2, where the commodity is unloaded. In order to move the loaded commodity from node n_1 to node n_2 the considered AGV has to travel through the lane network mentioned before following a path connection n_1 with n_2. If no AGV is waiting at node n_1 at the time where the task become known then an AGV has to travel empty (unloaded) from its current position n_0 in the network to n_1 (forward re-positioning). The repositioning task (n_0,n_1) has to be executed before the transportation task (n_1,n_2) can be executed. If the next transportation task assigned to the considered AGV requires a pickup operation at $n_3 \neq n_2$ then the AGV has to execute a forward re-positioning task (n_2,n_3) first. If the AGV has completed all assigned transportation tasks as well as all necessary forward re-positioning tasks then the AGV returns to a waiting area (backward re-positioning) where it remains idle until it is requested to fulfill further transportation requests. Thus, each trip of an AGV through the network starts with the fulfillment of a forward re-positioning task. Then, an alternating sequence of transportation tasks and forward re-positioning tasks is executed. At the end of a trip through the network, the AGV fulfills a backward-repositioning task.

This is the typical situation as known from pickup und delivery problems (PDP) investigated in the vehicle routing context. However, in AGV routing, there are several issues that let the AGV routing become a quite more complicated and even more challenging decision problem than the routing of manned-trucks in a road-network. The most important distinguishing fact is that due to the limited spatial extension of the terminal compared to a normal street map (used as network in PDP-scenarios) the AGV-trips are crossing or even partly overlapping with very high probability. In order to prevent accidents because two or more AGV are at the same position at the same time, it is necessary that the AGVs traveling in the network keep a security distance to each other. If the AGV-velocity is nearly constant then it is possible to ensure keeping of the security distance among the vehicles if the time between AGVs crossing the same node is at least M time units for all nodes (excepting the waiting area nodes).

To summarize the dispatching problem it is necessary to compile a least cost set of trips from the available transportation tasks for each AGV so that all transportation tasks are fulfilled and each AGV is assigned at most one trip. The generated trips have to be collision-free. A set of trips in a network is defined to be collision-save if for each node in the network the following property is fulfilled: There are at least M time units between two consecutively arriving AGVs. We refer to this decision problem as the static AGV-routing problem with collision-save trips. If the set of requests varies and the variations have to be reflected back into the trips of the AGVs then we refer to the corresponding decision situation as the dynamic AGV-routing problem with collision-free trips.

3. FORWARD-REPOSITIONING PLANNING PROBLEM

In this contribution, we restrict our investigations to a specific static AGV-routing scenario. We assume that some additionally appeared transportation tasks have been assigned to some AGVs. Now, these AGVs have to be set in duty, which means that they have to perform a forward-repositioning task from their current positions to the pickup location of the just received transportation task. Therefore, a collision-save path must be determined for each vehicle starting at the vehicle's current position, then visiting some node in the network and finally terminating in the node that is the pickup location of the assigned transportation task. The overall sum of travelled distances from each path should be minimized. We refer to this specific decision problem as the Forward-Repositioning-Planning-Problem (FRPP).

To prepare the application of computational decision support we formulate the FRPP as a mathematical optimization problem. Doing so, we interpret the (loaded or unloaded) AGVs as commodities for which secure routes (respecting the least distance condition) through the given terminal network are searched. Therefore, we propose a multi-commodity network flow model as adequate formalized problem representation.

Let G= (V, A, f) denote the (multiple-weighted) graph representing the network in which the AGVs are moving (cf. Fig. 1). The node set V contains all vertices in this network (including the current position of the AGVs outside regular operations areas) and the set A contains all arcs along which the AGV travel through the network. We use the multi-dimensional function f to describe the travel distance along the arcs in A as well as the travel times. If we assume that the AGVs travel with constant speed along the arcs then the least security distance is kept as long as in each node the arrival times of two consecutively arriving vehicles differ by at least M time units.

We assume that there is at most one request for each AGV. To fulfill the request it is necessary that the associated AGV travels from the corresponding start node to the associated stop node through the network represented by G. The selected AGV has to leave the start node within the associated time window (pickup time window) and must arrive at the destination node during the time window specified for the destination node (delivery time window). During the travel through the graph an AGV can wait at a node in order to postpone its arrival at another node.

We denote the set of available AGVs with K. Three families of decision variables are used to code the AGV-paths. The path-related **routing decisions** are stored in the binary decision variables x_{ijk} ($(i, j) \in A$, $k \in K$). If and only if AGV k travels along the arc $(i, j) \in A$ then $x_{ijk}=1$. Beside path-related routing decisions, we have to make **scheduling decisions** in order to determine the departure and arrival times of the AGVs from and at the nodes in the decided routes. Two families of continuous decision variables are declared to represent the schedule of the AGVs. The arrival time of vehicle k at node i is stored in at_{ik} and the departure time assigned to AGV k from node i is denoted by the decision variable dt_{ik}. We use these two continuous decision variables to check if the least distance M of preceding arrival times is kept. If the security distance restriction is considered in the flow optimization model then the flow optimization model becomes non-linear in general.

4. A DECODER-BASED HEURISTIC TO GENERATE COLLISION-FREE TRIPS

The previously described optimization model enables a simultaneous determination of the paths for the AGVs to be routed. However, if the model is a non-linear model with a quite large set of binary decision variables it is hardly to be expected that we can solve the model for instances of realistic seize. Therefore, we propose a heuristic solving approach that attempts to determine the needed AGV-paths consecutively. The basic idea of the heuristic is to determine a sequence $S=(s_1,\ldots,s_N)$ of the N considered AGVs. Such a sequence is then evaluated by trying to determine collision-free paths for the AGV in the order determined by S, i.e. S is evaluated by the costs resulting from following S in the path determination as well as by the number of AGVs for which we cannot generate a collision-secured paths. We use a genetic algorithm to improve the evaluated AGV-sequences.

The evaluation of S is as follows. At first, a shortest (cheapest) path is determined for the first AGV s_1 using the algorithm described in (Grünert and Irnich, 2005) for the shortest path problem with time windows (SPPTW). We select this path as collision-free trip for s_1 to perform the required forward-repositioning operation. According to this path, the AGV visits some nodes. Now, we fix an implicit security time window for each visited node. Within this security time window $[at_{ik}-M, dt_{ik}+M]$ it is not allowed that another AGV visits this node. The visiting time restrictions are added to the node description of node i and it must be respected by all subsequently routed AGVs. Now, we calculate the shortest path for AGV s_2 considering the added security time windows. If it is not possible to calculate such a path then we leave this vehicle unrouted and continue with AGV s_3. Otherwise, we select the generated path and add additional security time windows to the nodes visited by s_2 before we continue with the path determination for s_3.

After we have tried to find shortest paths for all N AGVs, we sum up the travel times and the number of AGVs for which we could not find a collision-secured path. Then we evaluate the sequence S by a rank-based fitness scheme in which the primary criterion is to maximize the number of routed AGVs and the secondary criterion is to minimize sum of travel times.

The incorporation of a construction heuristic into the genetic search space sampling of a permutation-evolving genetic algorithm is called a decoder-based approach (Michalewicz, 2000). Decoder-algorithms are used to manage complicated global constraints in optimization problems. In a computational simulation study, we proof the general applicability of this approach.

ACKNOWLEDGEMENTS
The authors are grateful to the generous support from the German Research Foundation and from the Korean Research Foundation.

5. REFERENCES

Grünert, T. and Irnich, S. (2005). *Optimierung im Transport, Band II: Rundreisen und Wege*. Shaker, Aachen, Germany.

Kim, K.H. and Bae, J.W. (2004). A Look-Ahead Dispatching Method for Automated Guided Vehicles in Automated Port Container Terminals. *Transportation Science, 38(2):* 224-234.

Michalewicz, Z. (2000). *Decoders.* Evolutionary Computation 2 – Advanced Algorithms and Operators. IoP, Brisol, United Kingdom

Möhring, R.H., Köhler, E., Gawrilow, E., Stenzel, B. (2005). Conflict-free real-time AGV Routing. *Proceedings of Operations Research 2004*, Springer, 18-24.

PLANNING FOR SELECTIVE REMARSHALING IN A CONTAINER TERMINAL USING EVOLUTIONARY ALGORITHMS

Kiyeok Park, Taejin Park, and Kwang Ryel Ryu*

Department of Computer Engineering
Pusan National University
Jangjeon-Dong, Geumjeong-Gu
Busan, Korea
e-mails: {nerissa79, parktj, krryu}@pusan.ac.kr

Abstract: Remarshaling in a container terminal refers to the task of rearranging containers stored in a stacking yard to improve the efficiency of subsequent loading onto vessels. Loading containers in the stacking yard, however, often cannot all be rearranged due to time constraints. This paper proposes a planning method for remarshaling loading containers in a selective manner. A crane schedule for rearranging selected containers is generated using a cooperative co-evolutionary algorithm that simultaneously searches for the best sequence of crane jobs and the best target configuration for stacking selected containers. Experimental results show that the proposed method outperforms other methods, which adopt other search algorithms.

1. INTRODUCTION

The efficiency of a container terminal heavily depends on the efficiency of loading containers onto vessels. The efficiency of loading is dependent on the configuration of containers stacked in the stacking yard, where the containers are temporarily stored, because loading begins by fetching containers from the stacking yard. Although containers are loaded onto vessels according to a predetermined schedule, if a target container is stored in the stacking yard while other containers are stacked above it, auxiliary operations, such as rehandlings that remove other containers, are required for a stacking crane to relocate those containers. With frequent rehandlings, loading can be delayed. In addition, if the stacking crane travels a long distance to pick a container, loading can also be delayed.

Loading delays caused by rehandlings or long travel distance can be avoided by simply respecting the loading schedule when piling containers. The loading schedule, however, is hardly known at the time when containers arrive at the stacking yard. Therefore ‘remarshaling’ emerged as an alternative method. It is a preparatory task that minimizes the delay of subsequent loading by arranging containers at the stacking yard during the idle time of stacking cranes.

Although there are many studies on remarshaling, we only focus on two: Park *et al*. (2009) and Park *et al*. (2010). Park *et al*. (2009) studied a planning method for remarshaling all containers using a cooperative coevolutionary algorithm (CCEA). The method divides the original planning problem into two sub-problems and tries to generate a better schedule using CCEA. Accordingly, compared to other algorithms, CCEA leads to better plans. However, Park *et al*. (2009) only focused on whether CCEA can lead to better plans as compared to other algorithms; they failed to consider that the idle time of stacking cranes is not enough to remarshal all containers. Park *et al*. (2010) focused on which containers should be remarshaled, rather than on how to remarshal containers. They proposed a selection method using a genetic algorithm (GA), and they evaluated the task scheduling using a heuristic method. In addition, to deal with estimation errors and uncertainty in the operation of stacking cranes, they proposed the iterative reselection method. Consequently, they reported that container selection can significantly affect the efficiency of loading and remarshaling.

In this paper, we propose a planning method for selective remarshaling using CCEA. In our proposed method, there are three sub-problems: the determination of a target configuration (Park *et al*., 2009), the determination of a moving priority (Park *et al*., 2009), and the selection of target containers (Park *et al*., 2010). We also adopt the iterative reselection method of Park *et al*. (2010) to improve scheduling quality. To assess the performance of the proposed method, we compare it with some other possible approaches. Experimental results show that, as compared to other approaches, CCEA can derive better schedules.

In the next chapter, we introduce some background information, such as the nature of the target container terminal and related works. In Chapter 3, a proposed method is described in detail, and the iterative reselection method is described in Chapter 4. Experimental results are reported in Chapter 5, and in Chapter 6, we finally give some concluding remarks.

2. BACKGROUND INFORMATION

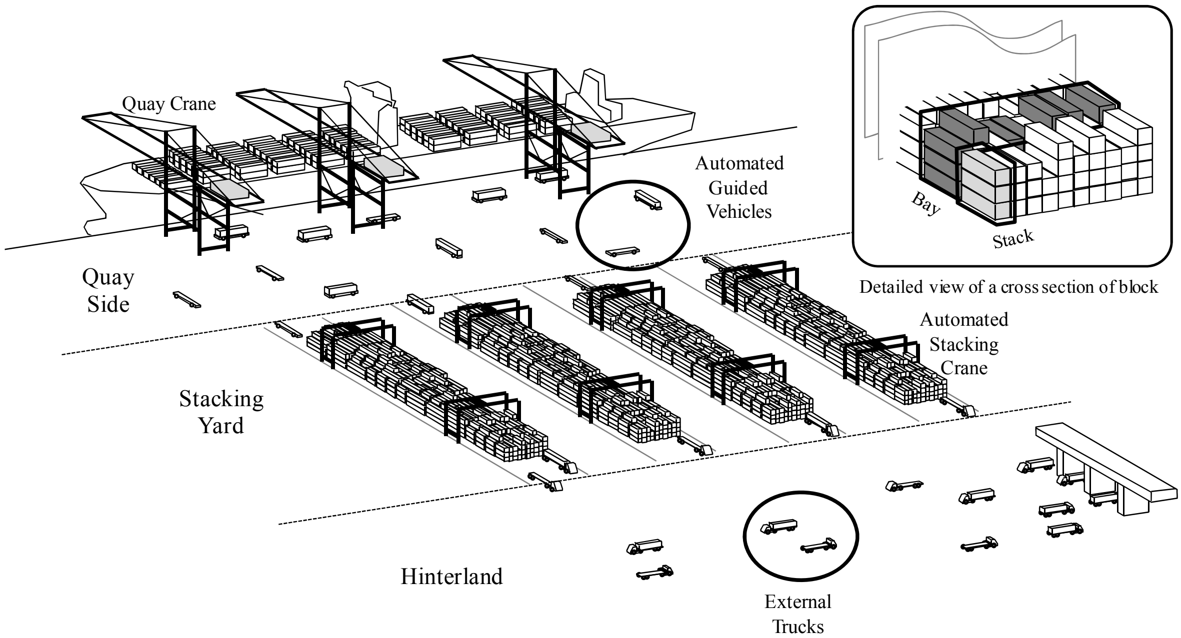


Figure 1. The layout of target container terminal

In this section, we give some background information to help readers understand the research. The container terminal, which is our target, is first described. Thereafter, the remarshaling of the target problem and some related works are introduced.

2.1 Container Terminal

Figure 1 illustrates the target container terminal of this study. The container terminal consists of three parts: the quay side, the stacking yard, and the hinterland. The quay side is where the vessel berths and where quay cranes (QCs) discharge containers from or load container onto vessels. The stacking yard, where containers are temporarily stored before import or export, is organized into several blocks, which consist of several stacks of containers. On each block, there are two automated stacking cranes (ASCs) that store or fetch containers. Automated guided vehicles (AGVs) deliver containers between these two types of cranes. In the hinterland, external trucks carry containers to and from the terminal.

Various equipment cooperate to process jobs in container terminals. For example, when loading starts, an ASC fetches a container from the stacking yard after such container was carried in by an external truck, and hands it over to an AGV. Subsequently, the AGV delivers the container to a QC. The QC ships the container, thereby finishing the loading operation. The points where external trucks or AGVs park to hand over containers to or to get containers from landside or seaside ASCs are called handover points (HPs).

2.2 Remarshaling in Container Terminals

The productivity of a container terminal is mainly dependent on the productivity of loading. If the ASC that fetches a container out from the stacking yard is delayed during loading, the delivery by AGVs or the shipping by QCs is also delayed. There are two main reasons that account for the delay of stacking cranes. First is the long travel distance of ASCs. Export containers are carried in by external trucks, which hand them over to the landside ASC at the landside HP. When the vessel arrives, the containers are fetched out by the seaside ASC, which hand them over to the AGV at the seaside HP. When export containers are carried into the terminal, they are stored near the landside HP to avoid interference between two ASCs. However, this results in a position that is completely opposite from the seaside HP where the containers are moved out to be loaded. Hence, the seaside ASC can be delayed because of the distance between the position where the containers are stored and the seaside HP.

The second reason is the occurrence of rehandlings during loading. Rehandling is an auxiliary job that retrieves other containers stacked above the target container. To avoid rehandlings during loading, the loading schedule is respected when loading containers. The loading schedule is determined considering two factors: the destination port of each container for convenience of operation at the destination port and the weights of each container to guarantee the safety of the vessel. When containers are carried into the terminal, however, the loading schedule is not yet determined. Although containers are located in positions, as determined by empirical rules, to prevent rehandling when containers are carried in, rehandling frequently occurs because of the estrangement between the configuration and the actual loading schedule.

To minimize the foregoing causes of delay, containers should be remarshaled before loading starts. There are several studies on planning for remarshaling. Kim *et al.* (1998) studied remarshaling in a conventional terminal where the stacking yard is laid horizontally against the quay side. They proposed the moving of containers to other blocks near the quay. Hirashima *et al.* (2006) proposed a method using reinforcement learning to derive a remarshaling schedule in the bay of a block. Kang *et al.* (2006) proposed a method for remarshaling in a conventional terminal using a partial order constraint graph and the simulated annealing algorithm. Lee *et al.* (2008) proposed the optimization model to improve the efficiency of remarshaling in a bay. Park *et al.* (2009) proposed a method using CCEA to derive a remarshaling plan in an automated container terminal where the stacking yard is positioned perpendicularly against the quay side. These studies focused on deriving remarshaling plans to minimize remarshaling time and to maximize loading efficiency—assuming that there is enough time to remarshal all containers. On the other hand, Park *et al.* (2010) focused on container selection to improve loading productivity when the given time for remarshaling is too short to remarshal all containers. They proposed a container selection method using a genetic algorithm and the iterative reselection method to cope with the uncertainty of ASC operation and the estimation error for remarshaling time. In addition, they also proposed a method to improve the efficiency of search by reusing the final population of the previous iteration as the initial population of the current iteration.

3. PLANNING FOR SELECTIVE REMARSHALING USING CCEA

Given that the amount of time for selective remarshaling is limited, planning for selective remarshaling prepares loading to enhance loading efficiency. Selective remarshaling focuses on minimizing delays within the given time. The suite of selected containers affects the performance of remarshaling in selective remarshaling; hence, containers that are located far from the seaside HP and that require much rehandling are selected within the given time. The actual remarshaling time, however, is hardly known until the position of each container is determined and remarshaling according to a schedule is simulated. In this study, therefore, we use a simulation to estimate the operation time after having determined the position of each container using CCEA.

3.1 Problem Decomposition and Representation for Each Sub-problem

The performance of CCEA depends on the decomposition of the original problem. There is a trade-off between fine-grained search and the maintenance of diversity. Numerous sub-problems can result in a narrow search space for algorithms, although the algorithms can suffer from the lack of population diversity corresponding to each sub-problem. Moreover, if there is strong interdependency among sub-problems, the stagnation of one sub-problem results in the stagnation of the other problems.

0	1	1	0	0	0	…	1	0

(a) The example of an individual for container selection problem

D_2	C_{10}	C_1	D_1	C_2	C_n	…	D_k	C_5

(b) The example of an individual for target configuration problem

C_{10}	C_1	C_3	C_7	C_{11}	C_9	…	C_n	C_4

(c) The example of an individual for moving priority problem

C_n : containers to be moved D_k : dummy containers

Figure 2. Individuals for each sub-problem

In this paper, we employ the decomposition used in Park *et al.* (2009), and we add a sub-problem for container selection. The given problem is therefore decomposed into two sub-problems: the problem of determining the target configuration and the problem of determining the moving priority. Subsequently, a population is assigned to each sub-problem to optimize each sub-problem independently. Individuals of each population, however, cooperate to evaluate an individual and to make a solution. The problem on target configuration determines where to store containers during remarshaling, and the problem on moving priority determines the priority to be used when deciding which container is to be moved first.

Figure 2 shows the individuals of each population. The individual for the container selection is represented as a binary string to imply whether a container is selected as a target or not. Its length is naturally the same as the number of loading containers. The individuals for the determination of the moving priority and of the target configuration are represented as permutations of containers. The position of containers in each individual represents the priority of containers. For example, in Figure 2(c), C_{10} is the most prior container, and C_4 is the least prior container.

The individual for the target configuration, however, does not mean a target configuration explicitly because a target configuration depends on an individual of the container selection problem. It is just a sequence of containers, and the real target configuration can be obtained after interpretation in the evaluation step. Making a solution follows choosing candidate positions which refer to as target slots where containers can be stored. If containers are selected, each container requires its own position for storage. To prepare for the assignment of a container to a position, we choose target slots using the method proposed by Park *et al.* (2009). The length of individuals is the same as the number of target slots. However, the length of individuals can be more than the number of containers. As such, to fill an individual up, we make some dummy containers. The dummies do not only fill an individual up but also help make various target configurations.

3.2 Creating and Evaluating a Candidate Solution

Individuals in each population cannot resolve a given sub-problem, and be evaluated individually. We should make an assessable schedule, including which containers are selected, where selected containers are stored, and what sequence is to be followed in moving containers by assembling individuals in each population. A schedule is evaluated by a terminal simulation system. The system is based on the specifications of a real container terminal, and it can simulate not only the status of a stacking yard but also the movement of the stacking crane, including acceleration and deceleration.

Figure 3 shows the derivation of a candidate solution from individuals gathered from each population. This derivation process is almost the same as the method proposed by Park *et al.* (2009), except that an individual container selection problem is used. The original method involves three steps: making a target configuration, deriving a partial-order, and deriving a remarshaling schedule. A target configuration is obtained by successively assigning selected or dummy containers to target slots according to the order found within individuals [Figure 3(a) and (b)], and by removing dummy containers and sorting the containers in each stack with respect to the loading schedule [Figure 3(c)]. After deriving the target configuration, we can get orders to store containers in each stack without rehandling. We obtain the ‘partial order’ [Figure 3(d)] because we only have the order in a stack not the order among stacks. Finally, we can derive a schedule from the partial order and the individual of moving priority.

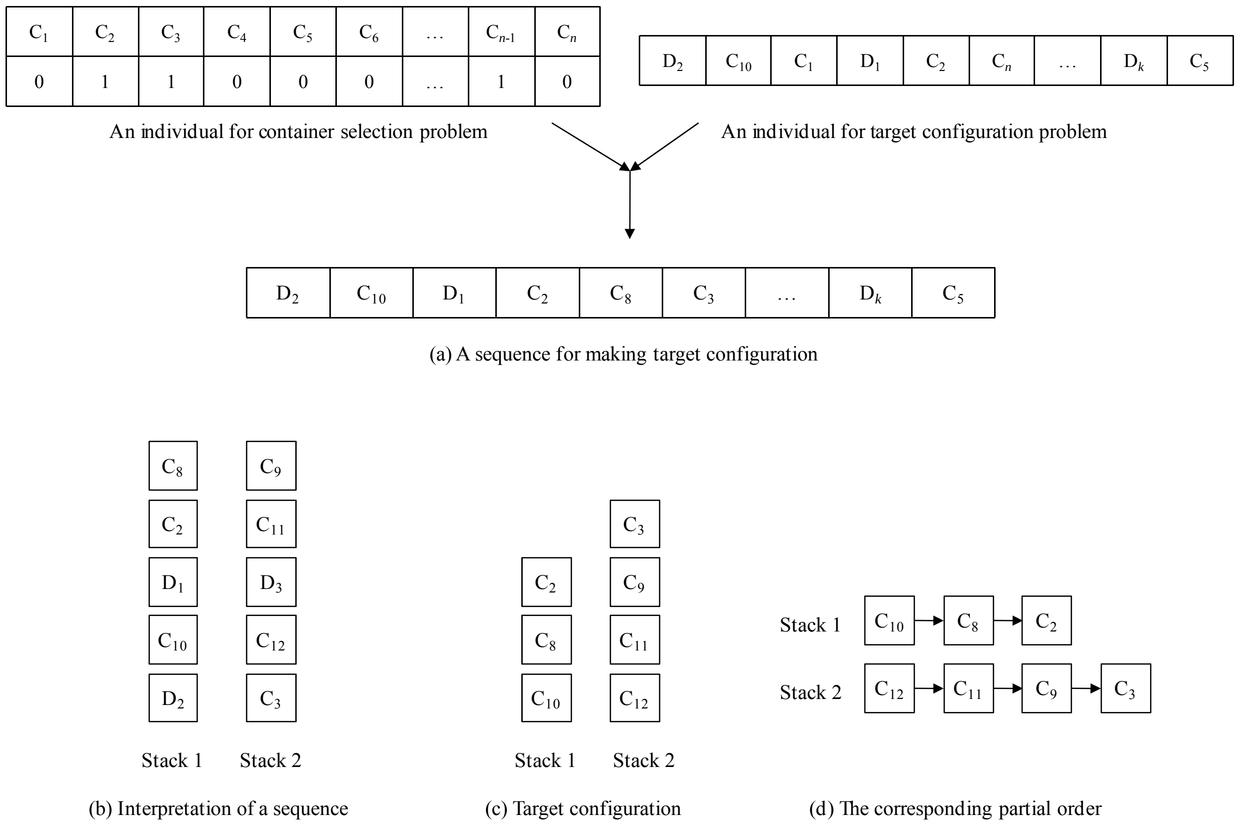


Figure 3. Example of creating a solution

4. ITERATIVE CONTAINER SELECTION

In the real world, the operation that obeys the schedule determined using the algorithm can be baffled with the uncertainty of movement of ASCs because estimated remarshaling time can differ from real operation time; that is, remarshaling cannot be done within a given time, or it can be terminated too early.

To minimize the gap between the estimated time and the real operation time, Park *et al.* (2010) proposed the iterative selection method, which varies from the iterative replanning method proposed by Ahn *et al.* (2007). This method reselects containers that will be remarshaled within the remaining time and replans with the reselected containers. Necessarily, containers that are already remarshaled are removed from the target container pool for reselection.

The reselection can be made, and with some advantages, using CCEA. A straightforward way of using CCEA is to apply it as if a new problem for selection is given in each iteration. However, because reselection in an iteration is not expected to make a dramatic change in the suite of containers selected in the previous iteration, the search for reselection can be made more efficient using the information obtained by the search conducted in the previous iteration. Instead of generating an initial population from scratch, the final population of the previous iteration is taken and adapted for reuse (Park *et al.*, 2010). The adaptation is conducted simply by eliminating the genes corresponding not only to the containers already processed in the previous iteration but also to those expected to be processed in the current iteration. By reusing the results of the previous search, a good solution is derived in less time.

5. EXPERIMENTAL RESULTS

5.1 Alternative Approaches for Comparison

Not only CCEA but also other approaches can be used to combine the two previous studies. The simplest one is selecting containers by GA (Park *et al.*, 2010), and then deriving a schedule by CCEA (Park *et al.*, 2009). However, because this approach shows the worst performance in experiments due to the estimation error of the heuristic used in evaluating the suite of containers, we exclude it from the set of approaches to be compared. Another approach, called GA in this section of the experimental results, is using a GA that acts like CCEA. It has three populations like the CCEA, but only one population evolves from a generation. It has already been used as an alternative approach in Park *et al.* (2009). The third approach, called NEST in this section of the experimental results, is using the GA used in Park *et al.* (2010), and then using CCEA of Park *et al.* (2009) instead of a heuristic when evaluating an individual.

5.2 Experimental Environment

The experiments have been conducted assuming an automated container terminal having the following conditions and features. A single vessel is berthing at the quay, and a thousand containers to be loaded onto the vessel are stored in seven blocks at the yard. Each block is equipped with two same-sized ASCs, and only the seaside ASC is used for remarshaling. Ten different datasets, generated based on the statistics collected from a real terminal, have been used for ten rounds of experiments, and the results have been averaged.

To measure improvements in loading efficiency, the average operation time per container of ASC at loading is measured through the simulation. Our goal is to minimize delay in the stacking yard at loading; hence, we do not simulate the quay side operation, which necessitates the simulation of vehicles. The uncertainty of movement of the stacking crane is set following the normal distribution.

Table 1 shows various algorithm parameters. To maintain the same number of evaluation, the population size is controlled across the algorithms. Binary tournament selection is used for parent selection for every algorithm. Order crossover (OX) and exchange mutation, which exchanges two random genes of an individual, is used as reproduce operator for determining the target configuration and movement priority. Uniform crossover and bit-wise mutation is used for the selection problem.

Table 1. Experimental Parameters

CCEA		GA	NEST
Population size	50 for each population	150	30 for selection 25 for each of others
No. of Evaluation	75,000		
No. of Generation	500 5	00	50[1]
Mutation Rate	1/*L* for selection (*L*: individual length) 0.1 for others(10% of individuals is mutated.)		

[1]NEST spends 1,500 evaluations (30×25×2).

5.3 Results

Figure 4 shows the average operation time per container of the seaside ASC at loading after remarshaling according to a schedule derived by each algorithm. The line at the top of the graph represents the operation time at loading without remarshaling. Results from every algorithm are below the operation time of the case without remarshaling, and the cracks become deeper with more given time. As can be seen, CCEA outperforms the other algorithms.

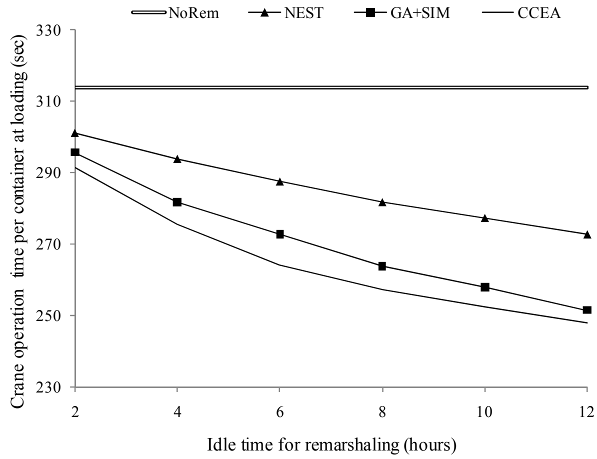


Figure 4. Average operation time per container of ASC at loading for each algorithm

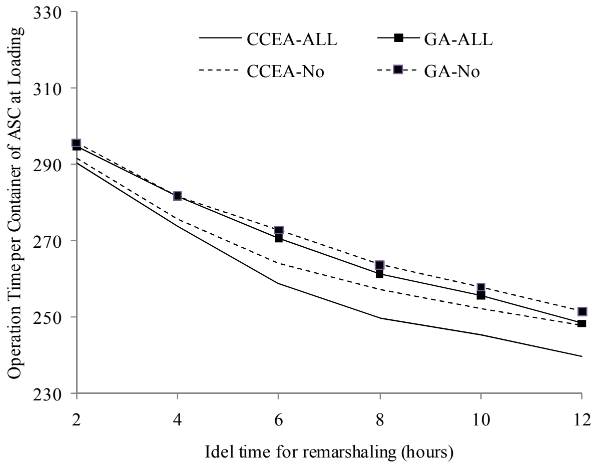


Figure 5. Average operation time per container of ASC at loading when previous population is reused

The su bstantiation fo r t his r esult i s pro vided i n Tabl e 2 . Tabl e 2 sh ows t he n umber of rem oved cont ainers (Remove) and remarshaled containers (Remarshal) during remarshaling, and the number of rehandlings during loading (RHatLoad) when 12 hours is given for remarshaling. CCEA can derive a sc hedule that remarshals more containers than the others, and after remarshaling, the occurrence of rehandling during loading can be minimized. GA and NEST process si milar n umber o f cont ainers, but i n t erms of t he occu rrence of re handling d uring l oading, GA outperforms NEST because NEST has a very short generation time to select good containers.

Figure 5 shows the results of the adaptation of reusing the previous population in the CCEA and GA. In the figure, '-All' means that the whole population is reused in the current iteration, and 'No' means that nothing is reused. In the case of GA, every strategy reports the same performance, but in the case of CCEA, there is enhanced performance when the whole population is reused because CCEA can satisfactorily maintain the diversity of the population.

Table 2. Number of Containers Moved during Remarshaling and Loading

Re	move	Remarshal	RHatLoad
CCEA 22	.6	132.3	2.8
GA 32	.3	122.6	8.4
NEST 33	.1	122.7	25.9

6. CONCLUSION

This paper has proposed planning methods using the cooperative coevolutionary algorithm for selective remarshaling in an automated container terminal, granting that the given time available for remarshaling is not long enough to rearrange all the containers to be loaded onto vessels. In the proposed method, the given problem is decomposed into three sub-problems, which are optimized by their own populations.

To compare the performance of the proposed method, some alternative approaches that combine the two previous studies are suggested. Experimental simulation shows that the proposed method outperforms the other approaches because CCEA optimizes each sub-problem independently using corresponding populations, although populations can use the proper information of other populations through interaction during evaluation. CCEA is advantageous in this iterative replanning framework because it can reuse the results of the search obtained in the previous iteration. By reusing the final population of the previous iteration as the initial population of the current search rather than starting from scratch, CCEA finds good solutions in less time.

ACKNOWLEDGEMENTS

This work was supported by the Grant of the Korean Ministry of Education, Science, and Technology (The Regional Core Research Program/Institute of Logistics Information Technology).

7. REFERENCES

Ahn, E. Y., Park, K., Kang, B., Ryu, K. R. (2007). Real Time Scheduling by Coordination for Optimizing Operations of Equipment in a Container Terminal. *Proceedings of ICTAI 2007*, vol. 1, pp. 44-48.

Hirashima, Y., Ishikawa, N., Takeda, K. (2006). A New Reinforcement Learning for Group-Based Marshalling Plan Considering Desired Layout of Containers in Port Terminals. *Proc. IEEE Conf. Networking, Sensing and control*, pp.670-675

Kang, J., Oh, M , S., Ahn, E . Y., Ryu, K. R ., Kim, K, H. (2006). Planning for Intra-block Remarshalling in a Container Terminal, *19th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, vol. 4031, pp. 1211-1220.

Kim, K. H., and Bae, J. W. (1998). Re-marshalling Export Containers in Port Containers Terminals. *Computers and Industrial Engineering*, 35(3-4), 655-658.

Lee, Y., and Hsu, N.Y. (2007). An Optimization Model for the Container Pre-marshalling Problem. *Computer and operations Research*, vol.34, pp.3395-658

Park, T., Kim, J., and Ryu, K. R. (2010). Iterative Replanning using Genetic Algorithms for Remarshaling in a Container Terminal. *Proceedings of AIA 2010*, CD, ISBN-(Book)978-0-88986-817-5.

Park, K., Park, T., and Ryu, K. R. (2009) . Planning for Remarshaling in an Automated Container Terminal using Cooperative Coevolutionary Algorithms. *Proceedings of SAC 2009*, 1098-1105.

Potter, M. A. and DeJong, K. A. (2000). Cooperative Coevolution: An Architecture for Evolving Coadapted Subcomponents. Evolutionary Computation, 8(1), MIT press, pp. 1-29.

Invited Speech

·Day2: Sep. 16 (Thu.)

·Time: 14:40 - 16:00

·Chair: Il-Kyeong Moon

·Room: Ballroom C, 5F

FLEXIBLE LOGISTICS NETWORK MODEL: EVOLUTIONARY APPROACH AND CASE STUDY

Mitsuo Gen

Fuzzy Logic Systems Institute
Iizuka, Japan
e-mail: gen@flsi.or.jp

Abstract: By distributing the global market and increasing competition for the market share, the fundamental key is the **productivity** and **logistics** in a manufacturing system in which here is **the optimization problem** for designing an effective logistics network system to increase the productivity and service quality. In this talk we will concentrate a decision making on supply chain management and logistics system into a seamless, productive and flexible process supported by the latest **computational intelligence**.

In this presentation the background of **the productivity and logistics in manufacturing system** will firstly introduced for understanding its four functional divisions: design, planning, manufacturing and distribution. The typical optimization problems in each division will be introduced briefly. In order to solve various real-world optimization problems within **interactive response time**, **the evolutionary approaches** based on Genetic Algorithm in computational intelligence will be surveyed from ERP packages, Simulation software and its applications. Many real world applications involve highly complex issues arising from the complexity of the problem structure and constraints. In addition, multiple objectives are to be handled simultaneously which often makes the problem intractable using traditional approaches.

As the case study, the flexible logistics network model will be introduced. In the traditional **multistage logistics network** sometimes causes some problems, such as too long delivery path, slow response speed *etc*. In this talk, we address **multistage flexible logistics network** model with nonadjacent structure, *i.e.* some non-neighboring echelons are connected with arcs (nonadjacent connecting arcs). In some practical cases, the nonadjacent connecting arcs make the logistics networks cost-effective and adaptable to changes in situation. On other hand, the existence of them makes the problem much more difficult. We formulate this problem as **location-allocation model**, and introduce **an effective hybrid genetic algorithm** to solve the optimization problem for flexible logistics network. As the second case study, we treat **the integrated decision making process** combining **flexible logistics location-allocation problem and lot sizing problem** for considering the two factors. This attempt makes the problem much more difficult to be solved. We formulate this problem as **a mixed integer programming model** and introduce an effective self-controlled hybrid genetic algorithm to solve this problem. Moreover, numerical analysis of case study is carried out to show the effectiveness of the approach proposed.

This talk will conclude with discussions of the evolutionary approach for treating various logistics network optimization problems for increasing the productivity in the manufacturing system including the distribution.

BIOGRAPHY

Mitsuo Gen is a visiting scientist at **Fuzzy Logic Systems Institute**, Japan since August 2009 after retiring Professor at Graduate School of Information, Production and Systems, Waseda University in July 2009. He received a PhD in Engineering from Kogakuin University, Japan in 1975 and second PhD in Informatics from Kyoto University, Japan in 2006. He was a Lecturer during 1974–1980, an Associate Professor during 1980–1987, a Professor during 1987–2003 at Ashikaga Institute of Technology, Japan. He was a Visiting Assoc. Prof. at the Univ. of Nebraska-Lincoln in 1981and a Visiting Prof. at Univ. of California at Berkeley in 1999.

His research interests include **Genetic Algorithms**, **Fuzzy Logic**, Neural Network and their applications to **Scheduling**, **Logistics Network**, and Reliability Design. He has authored 5 books, such as Genetic Algorithms and Engineering Design, 432pp, John Wiley & Sons, New York (1997), Genetic Algorithms and Engineering Optimization, 512 pp, John Wiley & Sons, New York (2000) with Dr. R. Cheng and Network Models and Optimization: Multiobjective Genetic Algorithms, 710pp, Springer, London (2008) with Dr. R. Cheng and Dr. L. Lin. He edited Intelligent and Evolutionary Systems, 240pp, *Studies in Computational Intelligence*, vol. 187, Springer-Verlag, Heidelberg, 2009 with others. He published Introduction to Evolutionary Algorithms, 418pp, Springer, London (July 2010) with Dr. Xinjie Yu.

He published more than 500 journal/conference papers, 5 books written and 3 books edited. These articles and books have been cited more than 6,000 times by researchers in the world. His papers published from the following international journals: *IEEE Transaction on Reliability*, *European J. of Operational Research*, *J. of Operational Research Society*, *OR Spectrum*, *Networks*, *Inter. J. of Systems Science*, *Computers in Industry*, *Computers & Industrial Engineering*, *Computers & Operations Research*, *Inter. J. of Production Research*, *Inter. J. of Production Economics*, *J. of Intelligent Manufacturing*, *Expert Systems with Applications*, *Computer Integrated Manufacturing*, *Fuzzy Sets and Systems,* and so on.

He invited a chapter on Evolutionary Algorithms and applications in Springer Handbook of Engineering Statistics edited by Dr. Hoang Pham, Springer, 2006 and two chapters on Evolutionary Algorithms with Dr. L. Lin in Wiley Encyclopedia of Computer Science and Engineering edited by Dr. Benjamin Wah, John Wiley & Sons, 2009 and Springer Handbook of Automation edited by Dr. Shimon Y. Nof, Springer, 2009 respectively. His **three recent research papers** based on GAs published in the last five years in *Computers & Industrial Engineering* journal included among **Top 10 Cited Papers** in March 2010:
http://www.elsevier.com/wps/find/journaldescription.cws_home/399/description#description

He supervised or advised more than **23 PhD degrees** from Waseda University, Ashikaga Institute of Technology, and several universities in Japan. He had more than **15 Visiting Scholars** for Post-Doctors or Visiting Faculties from China, Korea, Turkey and others.

He is a member of the editorial board of the following international journals: Editor-in-Chief: *Industrial Engineering and Management Systems*, http://kiie.org/iems/, Area Editor: Computational Intelligence of *Computers & Industrial Engineering,* http://www.elsevier.com/wps/find/journaleditorialboard.cws_home/399/editorialboard, Co-Editor: *OR Spectrum,* Editorial Board: *Fuzzy Optimization and Decision Making, Inter. J. of Manufacturing Technology and Management* and others.

TECHNIQUES FOR INSERTING ADDITIONAL CONTAINER TRAINS INTO EXISTING TIMETABLES

Erhan Kozan

Mathematical Sciences
Queensland University of Technology
GPO Box 2434, QLD 4001, Australia
e.kozan @qut.edu.au

Abstract: Techniques for scheduling additional train services are considered as is train scheduling involving general time window constraints, fixed operations, maintenance activities and periods of section unavailability. The inserting additional train is important because additional services must often be given access to the railway and subsequently integrated into current timetables. The problem therefore considers the competition for railway infrastructure between new services and existing services belonging to the same or different operators. The problem is characterised as a hybrid job shop scheduling problem with time window constraints. To solve this problem constructive algorithm and metaheuristic scheduling techniques that operate upon a disjunctive graph model of train operations are utilised. A software is developed to accomplish the following three functions: view train schedules in any railway line or network; manipulate the schedule by changing trains itinerary using simple drag and drop operations; and solve the scheduling and re-scheduling problems in a reasonable time period using the proposed approaches. A case study of Queensland Railway is provided and demonstrates the quality of solution that may be obtained by the software.

Keywords: container trains, train scheduling, job shops, metaheuristics.

Session I1: Empty Container

· **Day2: Sep. 16 (Thu.)**

· **Time: 16:20 - 17:40**

· **Chair: Michael G. H. Bell**

· **Room: Camellia, 5F**

A COMPUTER SIMULATION MODEL OF EMPTY CONTAINER MOVEMENT

Lee Loo Hay[1*], Chew Ek Peng[1], Long Yin[1], Luo Yi[1] and Shao Jijun[1]

[1]Department of Industrial & System Engineering
National University of Singapore
1 Engineering Drive 2
Singapore, Singapore
E-mail: iseleelh@nus.edu.sg

Abstract: This paper presents a computer simulation model of empty container repositioning operations for ocean container carriers. This simulation model simulates both the transport of laden containers and empty containers by liner services operated between different ports. Customer demand of containers and number of returned containers at various ports are generated from historical data. At different ports, both laden and empty containers are loaded onto and offloaded from vessels, and then transported between ports according to predefined service schedule. The decision model used for loading and unloading empty containers is separated from the main simulation model. Thus it is possible to compare the performance of different repositioning policies.

1. INTRODUCTION

The container shipping industry has enjoyed eye-popping growth over the past 40 years. As a result, there are tremendous opportunities and fierce competition in the container shipping industry especially with the rapid economic development in several Asia-Pacific economies including China and India. However, due to the severe international trade imbalances between different regions, liner operators today often face a challenge to effectively operate empty containers in a dynamic environment, which involves huge cost and a lot of inefficiencies. The problem of Empty Container Repositioning (ECR) thus attracts much attention from the academia.

To deal with the empty container repositioning problem, a variety of methods have been proposed by a number of researchers. The literature review on this topic is presented here with more than 20 papers. These papers generally do not sufficiently address the complicated problem face by liners in real world which is characterized by a lot of different decisions and huge problem size.

The real operations of ocean container carriers are very complex. The first issue in the complex operation is the shipping network operated by carriers, which often includes numerous ports and services. Most liner companies operate weekly or bi-weekly services, and thus typically deploy several vessels on one service. Another important issue is the uncertainty involved in the operations. The customer demand for laden containers varies with time, and the details of the demand, including type, quantity, origination and destination, are not known by the ocean carrier in advance of the booking requests from the customers. The demand is thus considered as uncertain when making empty container decisions. Similarly, the number of empty containers returned from previous customers is also uncertain. Liner companies often use demand and supply (returned container) forecasts in making ECR decisions.

The problems involved in ECR can be categorized into two distinct groups. Some problems are of a longer term or strategic nature. Examples of such problems are designing shipping networks, establishing hub-and-spoke structure, and determining composition of owned and leased containers. Some problems are of a medium term or tactical nature. The design of the empty container management strategy, that determines how to repositioning empty container within the shipping network, is one example of tactical problems. Some of these problems are of a shorter term or operational nature. Examples include the day-to-day management of empty container movements, on how to move extra empty containers in surplus ports to deficit locations, and which route to choose. These various problems involve complex and uncertain aspects which make them very challenging. It is often hard to develop good and effective solutions to these problems, and it is even harder to evaluate the performance of the solutions. Therefore a simulation would be a useful tool to estimate the performance measures associated with the above mentioned decision problems, and thus evaluate the effectiveness of proposed algorithms.

The operation of empty container repositioning generally include the following: global repositioning, which determines the number of containers to reposition to other geographical region, typically on long haul services; regional repositioning, that is to move container between different ports within a geographical region; container leasing; container storage (in port or depot); and new container procurement. These operations are interrelated and greatly affect one another. The problem, with respect to effectively making ECR decisions, is how to unify these operations in a plan intended to minimize the total cost of ECR through the shipping network.

The organization of this paper is as follows. In the next section we provide an overview of the literature related to Empty Container Repositioning, especially with simulation modeling approach. The following section presents a brief description of our simulation modeling procedure for ECR. Also, this section is concerned with how to evaluate different solutions to the ECR problem. This is followed by the last section which gives some concluding remarks.

2. LITERATURE REVIEW

For many years, the container shipping industry has dealt with empty containers repositioning quite extensively, due to the huge costs and resources attached to it. However, due to the complexity of the problem, ocean carriers and other transport operators often find it difficult to get effective and robust solutions. Furthermore, empty container repositioning often need to be considered jointly with other problem in the liner shipping business, and thus creating new problems which is more complicated. These problems may include the service network design with empty container repositioning considerations, the container fleet sizing and leasing-ownership composition, the design of a storage inventory network to balance demand and supply of empty containers, and the availability of empty container vessel slot capacity for least-cost repositioning. These interesting problems have attracted substantial scientific research during the last fifteen years (Theofanis & Boile, 2009).

Much attention has been focused on the empty equipment allocation and distribution problem, as well as the problem of balancing demand and supply between ports to meet future demands, by many researchers. Crainic et al. (1993) studied the problem of allocating empty containers in a land distribution and transportation system by introducing two dynamic deterministic formulations for the single and multi-commodity cases. Gendron and Crainic (1995) used a branch-and-bound algorithm, in which bounds are computed by a dual-ascent procedure, to solve the multi-commodity location problem with balancing requirements. Gendron and Crainic (1997) presented a parallel branch-and-bound algorithm for solving the multi-commodity location problem with balancing requirements. Shen and Khoong (1995) proposed a decision support system based on network optimization models and heuristics to solve a large-scale planning problem concerning the multi-period distribution of empty containers for a shipping company. Olivo et al. (2005) formulated the empty container management problem as a min cost flow problem whose arcs represent services routes, inventory links and decisions concerning the time and place to lease containers from external sources. Cheung and Chen (1998) developed a two-stage stochastic network model to determine the empty container repositioning and leasing decisions. Lam et al. (2007) developed an approximate dynamic programming approach considering both two-port two-voyage system and multiple-port multiple-voyage system in deriving operational strategies for the relocation of empty container. Liu et al. (2007) jointly considered container flow management and ship deployment problem. Feng et al. (2008) developed a two-stage model to deal with the ECR problem involving 17 services for intra-Asia transportation.

Some researchers have also studied the dynamic equipment allocation and reuse problem. Jula et al. (2006) studied the how to reduce congestion empty container movements in Los Angeles and Long Beach port area by optimizing the empty container reuse. Belmecheri et al. (2009) proposed a mathematical model to optimize empty containers reuse between regional consignees, shippers, and port terminal.

Another popular problem that many researchers have looked at is the empty balancing strategies within the context of a network design problem. Imai and Rivera (2001) addressed the problem of fleet size planning for refrigerated containers, which is the case when an extremely unbalanced trade represents one of the major investment decisions to be taken by liner operators. Ang et al. (2007) studied the sea cargo mix problem in international ocean container shipping industry which is closely linked to empty container repositioning. Shintani et al. (2007) considered the design of container liner shipping service networks by explicitly taking into account empty container repositioning, and showed that considering empty container repositioning in the network design problem provides a more insightful solution than the one not considering.

There are some specific topological structures in the shipping network which could be exploited to solve empty container repositioning problems. Li et al. (2004) showed that the two-critical point policy is optimal for the single port system in the finite horizon, and extended their findings to the discounted infinite horizon case and the average infinite-horizon case with truncated state space. Song (2005) showed that the optimal stationary policy is of threshold type in a continuous-review two-depot system. Song (2007) then analyzed the periodic-review shuttle system and finds the optimal stationary policy which has a simple threshold structure. Song and Earl (2008) showed that the optimal empty repositioning policy for a particular fleet size is of the threshold control type, and then derived the optimal threshold values and fleet-size. Song and Carter (2008) extended the empty vehicle redistribution problem to a hub-and-spoke transportation system, with random demands and stochastic transportation times. Song and Dong (2008) applied the threshold policy for empty container management in a cyclic shipping route problem and demonstrated that the threshold policy significantly outperforms the heuristic policies with simulation results.

Simulation is not a new methodology in the ocean transportation area. But major attention has been put on port operations instead of liner operations. There are relatively few notable simulation models for liner operations. Lai et al (1995) developed a simulation model of a shipping company's operational activities to aid the search for the best empty container allocation policy. Van Rensburg et al (2005) developed a general computer simulation model of ocean container carrier operations which simulates the transport of containers by container vessels. Henesey et al (2009) developed an agent based simulator for evaluating operational policies in the transhipment of containers. McLean and

Biles (2008) built a simulation model of the operation of a liner shipping network that considers multiple service routes and schedules, to evaluate the operational costs and performance associated with liner shipping, as well as the impact of individual service schedules on the overall system.

Simulation modeling techniques have not been extensively applied to solve ECR problems by researchers. This paper aims to simulate the container movement of container carriers, both laden and empty, so as to evaluate costs and other performance measures in ECR operations.

3. DESCRIPTION OF THE SIMULATION MODEL

Micro-City, an open-source simulation software developed in-house was used to build the simulation model. The model simulates the operation of liners, considering multiple services, multiple vessels, intermodal container movements, loading and unloading activities on multiple ports and customer demand at each port.

The simulation model mimics the operation of a set of container vessels travelling between different ports on specific service schedule. Once a container ship arrives at a port, inbound laden are offloaded and outbound laden container are uploaded on the vessel. Moreover, the empty containers repositioned to the port via this vessel are also offloaded, and empty containers bound for other ports on the route are loaded. The total time the vessel stay at the port is predetermined by its schedule, and we assume all handling activities are finished with this stay time period. The ship then sails out of the port to the next stop. The travailing times for containerships between stops are also predetermined in the schedule. Container ships in the simulation model operate continuously once deployed until the end of the planning period. Multiple service schedules are included in the model, and more than one ship of a different route can call at the same port at the same time.

The model can be divided into two main entities and three events. The details of the simulation model are described next.

3.1 Vessel Entity

Vessels are important entities in the simulation model. The carrier operates vessels to transport containers, both laden and empty between ports in the network. The vessel entity defined in our simulation model has several properties and fields associated with it. The first property is the capacity of the vessel. Each vessel is defined with specific space and weight capacities. (TEU and DWT) The current laden and empty container inventories on the vessel are tracked with 2 entry fields. The current stop of the vessel and the next stop on the schedule are also kept in respective fields. The available capacity of the vessel can thus be calculated by subtracting the current inventory from the total capacity.

Vessels defined in the model are put on repeating services, which are often formed by a cycle of voyages (origin-destination links) according to input service schedule. A typical service often follows a predetermined frequency of departures at each port, and the most common frequency is once per week. The number of vessels deployed on a specifc service can be calculated by dividing the total travel time for a complete cycle of the service, by the frequency of departures scheduled in the service, and the result is rounded up for any decimal points. For example, if the total travel time for a service is at 46 days, and the service is a weekly service, then 7 vessels are needed to operate the service, and these 7 vessels should be equally spaced in a 7 day interval. In our simulation model, we deploy all vessels according to input services at the initializing phase of the model.

All services are inputted into the model by the Service Route table. The table contains the estimated arrival time and estimated departure time for the service at different ports in the route. It also includes information like the stay time at each port in the route and the travel time between consecutive stops.

3.2 Port Entity

Ports are also defined as entities in the model. Ports are the end nodes of each voyage. Every vessel starts and ends its sailing at a port. The container inventory at each port is divided into three subcategories in our simulation model. The first category is the available empty container inventory at the port. This is inventory that can be used directly by vessels to satisfy loading requests from customer demands. It is also available for empty container repositioning considerations. Another subcategory of container inventory is the newly unloaded laden containers. These containers are unloaded from vessels visiting the port and are not available for use immediately. In reality, these unloaded containers will be picked up by customers. In a week or two, the content of these containers would be cleared and they were added to available empty container inventory. The last subcategory of container inventory in the model is defined as dummy inventory holding the unloaded laden containers which have been in the port for one week and are not yet cleared as available empties. In the next week, these containers would be cleared and transferred to available empty container inventory.

At the port, laden containers are unloaded first, followed by any empty containers repositioned. Then the laden and empty containers are loaded onto the vessel.

3.3 Unloading Event

After vessels arrival at a port, the unloading event will trigger. In the unloading event, both laden and empty containers are offloaded from the vessel following the flow chart below.

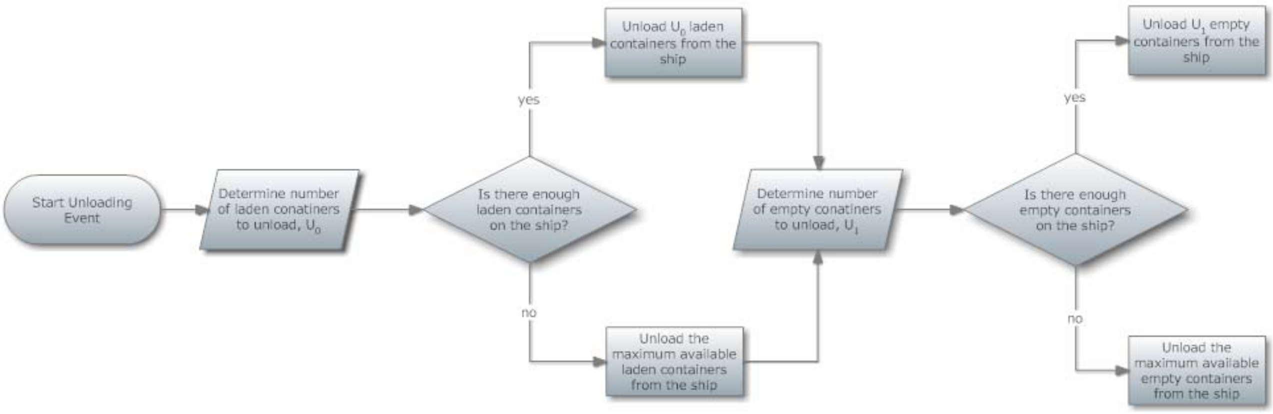


Figure 1. The Unloading Event

The number of laden containers to be unloaded is determined by the inbound containers at the port, which is generated from input parameters. In the current model, the inbound containers are assumed to follow normal distribution. It is also possible to change to other general distributions. The model will then check the number of laden container inventory on the ship to make sure there is enough to unload. If there is not enough laden containers to unload, the maximum number of laden containers on the ship will be unloaded to the newly unloaded laden container inventory of the port.

After unloading the laden container, the number of empty container to unload is also determined. This information is obtained from the empty container repositioning algorithm, which is a separate module possibly a standalone program. Similar to the laden unloading process, the availability check for empty containers is also performed and appropriate number of empty containers is unloaded from the ship. These unloaded containers go directly into the available empty container inventory of the port, and can be used by other vessels immediately.

At the end of the unloading event, a loading event for the same vessel is scheduled, considering the stay time of the vessel at the port.

3.4 Loading Event

The loading event for any vessel happens after its respective unloading event at the same port. In the loading event, both laden and empty containers are loaded to the vessel from the port, and the port's container inventory is also changed. The following diagram shows the logic flow of the loading event.

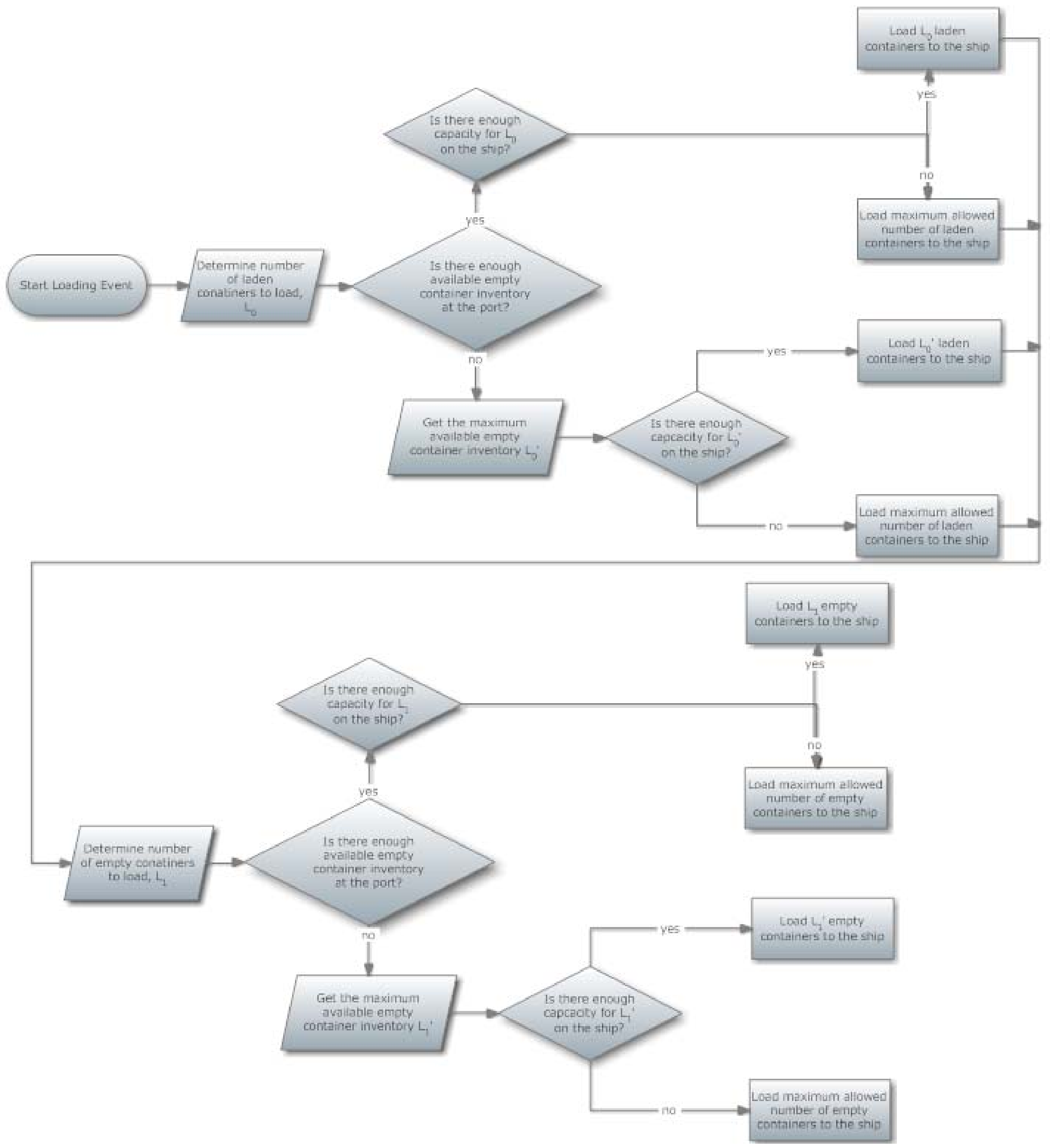


Figure 2. The Loading Event

Similar to the unloading event, the number of laden containers to be loaded is determined by the outbound demand at the port, which is also generated using similar methods to the inbound information. The model will then check the number of available empty container inventory at the port to make sure there is enough to load. Vessel capacity is also check to finally determine the amount of laden containers to load.

After loading the laden container, the number of empty container to load is determined with the same module in the unloading event. Similar to the laden unloading process, the availability checks on port's available empty container inventory and ship's capacity for empty containers are performed and appropriate number of empty containers is loaded to the ship.

At the end of the loading event, the next loading event for the same vessel is scheduled, considering the travel time from the current port to the next port on the route.

3.5 Updating Event

Besides the loading and unloading event, there is another important event in the simulation model: the updating event. The updating event is triggered at the beginning of each week and is in charge of updating the inventory at each port and generating inbound and outbound containers, as well as generating forecasting inbound and outbound information which is used in the empty container reposition module.

The updating event updates the three categories of container inventory at all ports following the formula below:

$$I_t = I_{t-1} + \alpha R_{t-1} + I^{r}_{t-1}$$

$$I^{r}_t = (1 - \alpha) R_{t-1}$$

Where I_t is the available empty container at week t, R_t is the unloaded laden containers during week t and I^{r}_t is the dummy inventory.

The inbound and outbound container at each port is also generated in the updating event. In the current model, the information is generated using historical data of the past inbound and outbound values at all the ports. The updating event will also invoke the empty container repositioning algorithm to get the repositioning decisions for all vessels for the following week.

3.6 The Empty Container Repositioning Module

One of the goals of the simulation model is to compare the performance of different empty container repositioning algorithms. In the simulation model, there is a separated module which represents the empty container repositioning algorithms currently running. For example, we have developed a mathematical model for operational empty container decision making. The model is coded in C++ and uses Cplex solver. The simulation model will just make a call to the C++ executable file and read the output from the C++ program.

4. FUTURE WORKS

This simulation model is an ongoing development. We have planed for the addition of a user interface for the simulation to improve the interaction between the user and the simulation, some additional ECR algorithm modules for performance comparison, better demand and supply models, and models of delays and disruptions. It is also possible to make this model into a training tool which can be used to train industry decision makers.

ACKNOWLEDGEMENTS

The authors want to thank Naptune Orient Line and Maritime Authority of Singapore for their support and help in making this study possible. We also want to thank Mr. Arpan and Ms. Selina for their valuable input in the operation of empty container repositioning.

5. REFERENCES

Ang, J., Cao, C., & Ye, H.-Q. (2007). Model and algorithms for multi-period sea cargo mix problem. *European Journal of Operational Research , 180* (3), 1381-1393.

Belmecheri, F., Cagniart, T., Amodeo, L., Yalaoui, F., & Prins, C. (2009). Modelling and optimization of empty container reuse: A real case study. *International Conference on Computers & Industrial Engineering*, (pp. 1106 - 1109).

Cheung, R., & Chen, C.-Y. (1998). A two-stage stochastic network model and solution methods for the dynamic empty container allocation problem. *Transportation Science , 32* (2), 142-162.

Choong, S., Cole, M., & Kutanoglu, E. (2002). Empty container management for intermodal transportation networks. *Transportation Research Part E: Logistics and Transportation Review , 38* (6), 423-438.

Crainic, T. G., Gendreau, M., & Dejax, P. (1993). Dynamic and stochastic models for the allocation of empty containers. *Operations Research , 41* (1), 102-126.

Feng, C.-M., & Chang, C.-H. (2008). Empty container reposition planning for intra-Asia liner shipping. *Maritime Policy and Management , 35* (5), 469-489.

Gendron, B., & Crainic, T. (1995). A branch-and-bound algorithm for depot location and container fleet management. *Location Science , 3* (1), 39-53.

Gendron, B., & Crainic, T. (1997). A parallel branch-and-bound algorithm for multicommodity location with balancing requirements. *Computers and Operations Research , 24* (9), 829-847.

Henesey, L., Davidsson, P., & Persson, J. A. (2009). Agent Based Simulation Architecture For Evaluating Operational Policies In Transshipping Containers. *Autonomous Agents and Multi-Agent Systems , 18* (2), 220-238.

Imai, A., & Rivera, F. (2001). Strategic fleet size planning for maritime refrigerated containers. *Maritime Policy and Management , 28* (4), 361-374.

Jula, H., Chassiakos, A., & Ioannou, P. (2006). Port dynamic empty container reuse. *Transportation Research Part E: Logistics and Transportation Review , 42* (1), 43-60.

Lai, K., Lam, K., & Chan, W. (1995). Shipping Container Logistics and Allocation. *Journal of the Operational Research Society , 46* (6), 687-697.

Lam, S.-W., Lee, L.-H., & Tang, L.-C. (2007). An approximate dynamic programming approach for the empty container allocation problem. *Transportation Research Part C , 15* (4), 265-277.

Li, J.-A., Liu, K., Leung, S. C., & Lai, K. K. (2004). Empty container management in a port with long-run average criterion. *Mathematical and Computer Modelling , 40*, 85-100.

Liu, X., Ye, H.-Q., & Yuan, X.-M. (2007). *A Tactical Planning Model for Liner Shipping Companies: Managing Container Flow and Ship Deployment Jointly.* National University of Singapore, School of Business.

McLean, A., & Biles, W. (2008). A Simulation Approach To The Evaluation Of Operational Costs And Performance In Liner Shipping Operations. *Proceedings of the 2008 Winter Simulation Conference*, (pp. 2577-2584).

Olivo, A., Zuddas, P., Di Francesco, M., & Manca, A. (2005). An operational model for empty container management. *Maritime Economics and Logistics , 7* (3), 199-222.

Shen, W., & Khoong, C. (1995). A DSS for empty container distribution planning. *Decision Support Systems , 15* (1), 75-82.

Shintani, K., Imai, A., Nishimura, E., & Papadimitri, S. (2007). The container shipping network design problem with empty container repositioning. *Transportation Research Part E , 43* (1), 39-59.

Song, D. P. (2007). Characterizing Optimal Empty Container Reposition Policy in Periodic-Review Shuttle Service Systems. *Journal of the operational research society , 58*, 122-133.

Song, D. P. (2005). Optimal Threshold Control of Empty Vehicle Redistribution in Two Depot Service Systems. *IEEE Transaction on Automatic Control , 50*, 87-90.

Song, D., & Carter, J. (2008). Optimal Empty Vehicle Redistribution for Hub-and-Spoke Transportation Systems. *Naval Research Logistics , 55*, 156-171.

Song, D., & Dong, J. (2008). Empty Container management in Cyclic Shipping Routes. *Maritime Economics & Logistics , 10*, 335-361.

Song, D., & Earl, C. (2008). Optimal empty vehicle repositioning and fleet-sizing for two-depot service systems. *European Journal of Operational Research , 185* (2), 760-777.

Theofanis, S., & Boile, M. (2009). Empty marine container logistics: facts, issues. *GeoJournal , 74*, 51-65.

van Rensburg, J. J., He, Y., & Kleywegt, A. J. (2005). A Computer Simulation Model Of Container Movement By Sea. *Proceedings of the 2005 Winter Simulation Conference*, (pp. 1559-1566).

A FREQUENCY-BASED MARITIME CONTAINER ASSIGNMENT MODEL

Michael G H Bell, Xin Liu, Panagiotis Angeloudis, Achille Fonzone, and Solmaz Haji Hosseinloo

Port Operations Research and Technology Centre
Department of Civil Engineering
Imperial College London
London SW7 2BU
Email: m.g.h.bell@ic.ac.uk

Abstract: A novel global maritime container assignment model, based on the frequency-based transit assignment method of Spiess and Florian, is presented. In this model, containers are carried by shipping lines operating strings (or port rotations) with given service frequencies. An origin-destination matrix of full containers is assigned to these strings to minimize sailing time plus container dwell time at the origin port and any intermediate transhipment ports. This necessitated two significant model extensions. The first involves the repositioning of empty containers so that a net outflow of full containers from any port is balanced by a net inflow of empty containers, and vice versa. As with full containers, empty containers are repositioned to minimize the sum of sailing and dwell time, with a facility to discount the dwell time of empty containers in recognition of the absence of inventory. The second involves the inclusion of an upper limit to the maximum number of container moves per unit time at any port. The dual variable for this constraint provides a shadow price, or surcharge, for loading or unloading a container at a congested port. Insight into the interpretation of the dual variables is given. Model behaviour is illustrated by a numerical example. The paper concludes by considering the next steps toward realising a container assignment model that can, amongst other things, support the assessment of supply chain vulnerability to maritime disruptions.

A STUDY ON STOCHASTIC EMPTY CONTAINER REPOSITIONING PROBLEM

Long Yin, Lee Loo Hay*, and Chew Ek Peng

Department of Industrial and Systems Engineering
National University of Singapore
10 Kent Ridge Crescent, 119260, Singapore
*Correspondence: iseleelh@nus.edu.sg

Abstract: Severe international trade imbalances exist in different regions. One of the challenges faced by liner operators today is to effectively operate empty containers in order to meet demand and to reduce inefficiency in the uncertain environment. We formulate a two-stage stochastic programming model which incorporates random demand, supply and ship capacity, where the objective is to minimize expected total operational cost of managing empty container flow. Random sample scenarios are generated and the expected value function is approximated. Numerical experiments are also provided.

1. INTRODUCTION

Containerization has become more and more important in international freight transportation since 1970. In 2004, over 60% of the world's maritime cargo is transported in containers, while some routes among economically strong countries are containerized up to 100% (D. Steenken et al., 2004). One main issue in containerization is the imbalance of container flow because of imbalanced global trade between different regions. Thus, effectively repositioning empty container from surplus regions to deficit regions becomes more and more important. It is reported that empty containers have accounted for at least 20% of global handling activity ever since 1998 (Drewry Shipping Consultants, 2006/07). And it is forecasted that the total cost of empty container repositioning (ECR) will increase from US$1.98 billion in 2006 to US$2.58 billion in 2010 (Feng and Chang, 2008).

There are increasing studies on empty container management in recent years (Erera et al., 2005; Lam et al., 2007; Feng and Chang, 2007; Chang et al. 2008; Zhou and Lee, 2008; Imai et al., 2009; Zhang et al., 2009; Moon et al., 2010). There are several studies taking into account the uncertain nature in recent years (Cheung and Chen, 1998; Dong and Song, 2009). Erera et al. (2007) developed a robust optimization framework based on time-space network. The robust repositioning plan was developed based on the nominal forecast value and could be recovered under a set of recovery sections. Di Francesco et al. (2009) proposed a multi-scenario model, where shipping companies' opinions were considered to generate scenarios when the distribution of uncertain parameter cannot be estimated through historical data.

This paper is motivated by a shipping company which provides international maritime transportation service. Shipping companies usually forecast demands and supplies based on the historical inbound and outbound container, balance trend, delinquency record and economic environment in the future. In this case the approximate distribution of some uncertain parameters could be estimated based on this information. We develop a stochastic model to make short-term decisions taking account into some uncertain parameters. Scenarios are generated based on the estimated distribution of the uncertain parameters. The Sample Average Approximation (SAA) method is applied to solve the stochastic programs.

2. PROBLEM FORMULATION

One of challenges faced by the container operators is the uncertain characteristics in container transportation. Currently, container operators make decisions based on the expected parameters, which may cause inefficient in ECR. Because of the difference between the forecasted and realized value, infeasible solutions may be provided and have to be adjusted at real time operation. In this study, we develop a stochastic programming model which considers three main uncertain parameters: the demand (the container that picked up by the customers to load cargos), the supply (the empty container that returned by the customer), and the ship available capacity for empty containers.

This study aims to make short-term decisions that usually within one or two weeks and we also consider the influence of these decisions in following several weeks. The transportation network we considered is decomposed according to the geographical region and each region is considered separately. This network could be divided into two stages. We assume that all information in the first stage is known and we know the distribution of the uncertain parameters in the second stage. The deterministic formulation with single scenario could refer to Long et al. (2010), which considers the demand and supply of empty containers at each port, the liner services, the ship residual capacity and characteristics at different ports. Our model is run in a rolling horizon manner. ECR decisions are made at the beginning of stage 1 and will be made again when new information is collected.

Let $\omega \in \Omega$ denotes a scenario that is unknown when the first stage decisions are made, but that is known when the second stage decisions are made, where Ω is the set of all scenarios. In this ECR problem, we use $\{D(\omega), S(\omega), E(\omega)\}$ to represent the realized demand, supply and residual ship capacity respectively. A typical two-stage model for ECR is explained as follow:

Stage 1:

$$\min \quad g(x_1) = c_1 x_1 + E_p[Q(x_1, \xi(\omega))] \tag{1}$$

$$s.t. \quad A_1 x_1 = a_1 \tag{2}$$

$$B_1 x_1 = v \tag{3}$$

$$x_1 \geq 0 \tag{4}$$

x_1: Decisions at stage 1
c_1: The cost vector at stage 1
A_1, B_1: The coefficient matrices of x_1 in constraints (2) and (3) respectively
a_1: The RHS of constraint (2) at stage 1
v: The vector of end container status of stage 1. The end container status includes the empty container inventory at each port at the end of stage 1, and the number of empty container at each vessel at the end of stage 1. Note that v is also the initial container status of stage 2. The two stages are connected by v.

The objective function (1) is to minimize the total operational cost in the planning horizon. $c_1 x_1$ is the cost at stage 1, $E_p[Q(x_1, \xi(\omega))]$ is the expected cost at stage 2. We assume that the probability distribution p on Ω is known in the stage 1. Constraint (2) includes typical constraints of ECR problem (ship capacity constraint, service flow constraint, port flow constraint, across region flow constraint). Constraint (3) is to set container status of at the end of stage 1.

Stage 2: For a realized scenario ω, we have

$$Q[x_1, \xi(\omega)] = \min \quad c_2 x_2(\omega) \tag{5}$$

$$s.t. \quad A_2 x_2(\omega) = a_2(\omega) \tag{6}$$

$$B_2 x_2(\omega) = v \tag{7}$$

$$x_2(\omega) \geq 0 \tag{8}$$

$x_2(\omega)$: Decisions for scenario ω at stage 2
c_2: The cost vector at stage 2
A_2, B_2: The efficient matrices of x_2 in constraints (5) and (6) at stage 2
$a_2(\omega)$: The RHS of constraint (2) in stage 2. It is related to the stochastic variable $D(\omega), S(\omega), E(\omega)$.
Constraint (7) is to set the initial container status of stage 2.

3. METHODOLOGY

Our stochastic problem is difficult to solve because it is difficult to evaluate the expected cost of the second stage $E_p[Q(x_1, \xi(\omega))]$ for a given x_1. It requires the solution of a large number of second-stage optimization problems. In this study, we apply SAA method in addressing the large set Ω. The basic idea of SAA method is that the expected objective function of the stochastic problem is approximated by a sample average estimate derived from a random sample and the resulting SAA problem could then be solved by deterministic optimization techniques (Kleywegt, et al., 2002). Several replications with different samples are run to obtained obtain candidate solutions. This method has been applied to solve two-stage stochastic routing problems by Verweij et al. (2003).

In sample methods, a sample of ω^1, ω^2, …, ω^N of N sample scenarios is generated according to probability distribution p,

$$\hat{g}_N = \min \quad c_1 x_1 + \frac{1}{N}\sum_{n=1}^{N}[c_2 x_2(\omega^n)] \tag{9}$$

$$s.t. \quad (2), (3), (6), (7)$$

$$v \geq 0, x_1 \geq 0, x_2(\omega^n) \geq 0 \qquad \textit{for } n=1,2,\dots,N \tag{10}$$

The SAA problem shown is (9)-(10) then is solved using a deterministic optimization method. According to Nokin et al. (1998), Mak et al. (1999) and Verweij et al. (2003), the SAA method is introduced as follows. The SAA method proceeds by generating *M* independent samples with each sample size *N*, and solving the associated SAA problems. Then the objective values $\hat{g}_N^1, \hat{g}_N^2, \ldots, \hat{g}_N^M$ and candidate solution $\hat{x}_N^1$, $\hat{x}_N^2$, …, $\hat{x}_N^1$ could be obtained. The average of the optimal objective function values of the *M* SAA problem is denotes as $\bar{g}_N^M$, which provides a statistical estimate for a lower bound on the optimal value of the true problem (1)-(8). *N'* scenarios (N' >N) are then generated to evaluate the candidate solutions. $\hat{g}_{N'}$ is defined to estimate the objective value $g(\hat{x}_N)$ of an optimal solution $\hat{x}_N$ of the SAA problem, and it is an estimate of the upper bound on the optimal value of the true problem. It is natural to take $\hat{x}^*$ as one of the optimal solutions of the *M* SAA problems which has the smallest estimated objective value. The optimality gap could be estimated as $\hat{g}_{N'}(\hat{x}^*)$-$\bar{g}_N^M$, which can be used to evaluate the quality of the solution. The estimated variance of the gap estimator is defined as $\hat{\sigma}^2_{\hat{g}_{n'}(\hat{x}^*)-\bar{g}_N^M}$.

4. NUMERICAL STUDY

To evaluate the performance of the SAA method, we consider a network as shown in Figure 1. We consider 5 ports in the target region. The objective is to efficiently meet demands in the target region. *S1* and *S2* are across region services which travel between region A and region B. *S3* is an intra region service which travels between port 1 and port 5. Weekly service is considered and the schedule of each service is given and fixed. The planning horizon is three weeks. The first stage is the first week. The second stage is the second and third week. The lead time of across region empty container is one week. At the beginning of the first stage, the in-transit container in *S1* and *S2* which will arrive at the target region *A* in first stage is given. The across region empties which will arrive at region A at week 2 and week 3 in *S1* and *S2* is ordered at week 1 and week 2 respectively.

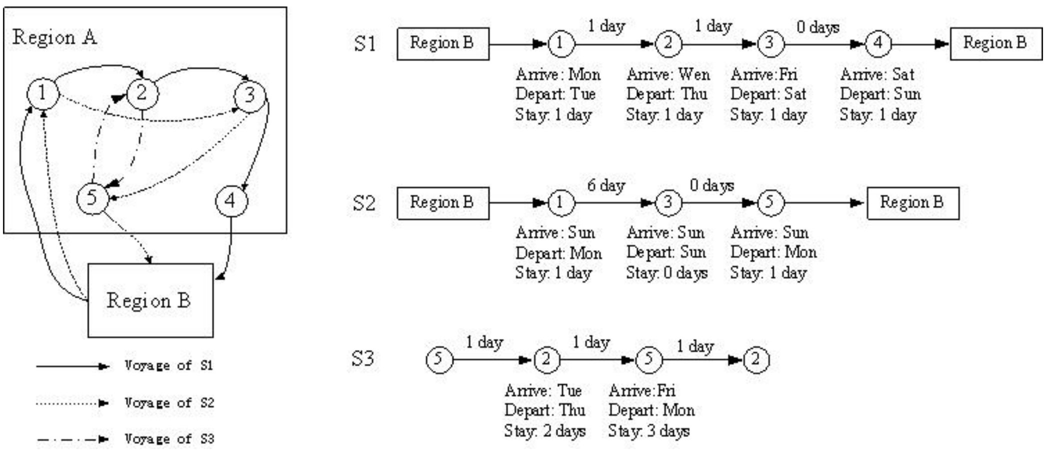


Figure 1. A network with three services and five ports

Replication number *M* is set to be 20. The sample size *N* is set to be 100. The number of scenarios to evaluate the solution is *N'*=1000. The performance of the SAA method is shown in Table 1. The lower bound on the optimal value of the true problem is 3006, and the upper bound is 3026.98. The results show that the optimal gap is 20.52 (0.68%) and the estimated variance of optimal gap estimator is 700.8. It indicates that we can get good solution by using SAA method (with *N*=100) to solve the stochastic ECR problem.

Table 1. Performance of the SAA method

$\bar{g}_N^M$	3006.46
$\hat{g}_{N'}(\hat{x}^*)$	3026.98
$\hat{g}_{N'}(\hat{x}^*)$-$\bar{g}_N^M$	20.52(0.68%)
$\hat{\sigma}^2_{\hat{g}_{n'}(\hat{x}^*)-\bar{g}_N^M}$	700.8

5. CONCLUSION

In this study, we develop a two-stage stochastic model to solve the ECR problem with uncertain demand, supply and ship residual capacity. SAA method is applied. The SAA problem with multi-scenarios for the real scale maritime transportation is usually has large scale and thus difficult to solve directly. It is shown by the numerical study that the optimal gap is quite small and we can get good solution.

6. REFREENCE

Annual Container Market review & Forecast 2006/07, Drewry Shipping consultants LTD: London.

Chang, H., Jula, H., Chassiakos, A. and Loannou, P. (2008). A Heuristic Solution for the Empty Container Substitution Problem. *Transportation Research Part E: Logistics and Transportation Review*, 44(2): 203-216.

Cheung, R.K. and Chen, C.-Y. (1998). A Two-Stage Stochastic Network Model and Solution Method for the Dynamic Empty Container Allocation Problem. *Transportation Science,* 32 (2): 142-162.

Di Francesco, M., Crainic, T.G. and Zuddas, P. (2009). The Effect of Multi-scenario Policies on Empty Container Repositioning. *Transportation Research Part E: Logistics and Transportation Review*, 45(5):758-770.

Dong, J.X. and Song, D.P. (2009). Quantifying the Impact of Inland Transport Times on Container Fleet Sizing in Liner Shipping Services with Uncertainties. *OR Spectrum*, 1-26.

Erera, A. L., Morales, J. C. and Savelsbergh, M. (2005). Global Intermodal Tank Container Management for the Chemical Industry. *Transportation Research Part E,* 41(6): 551-566.

Feng, C.M. and Chang, C.H. (2008). Empty Container Reposition Planning for intra-Asia Liner Shipping. *Maritime Policy and Management*, 35(5): 469-489.

Imai, A., Shintani, K. and Papadimitriou, S. (2009). Multi-port vs. Hub-and-Spoke Port Calls by Containerships. *Transportation Research Part E: Logistics and Transportation Review*, 45(5): 740-757.

Kleywegt, A.J., Shapiro, A. and Homem-De-Mello, T. (2002). The Sample Average Approximation Method for Stochastic Discrete Optimization. *SIAM Journal on Optimization*, 12(2): 479-502.

Lam, S.-W., Lee L.-H. and Tang L.-C. (2007). An Approximate Dynamic Programming Approach for the Empty Container Allocation Problem. *Transportation Research Part C,* 15(4): 265-277.

Long, Y., Lee, L.H., Chew, E.P., Luo, Y., Senguta, A., and Chua, M.S. (2010). A Time Space Network Model on Empty Container Flow Management. Submitted paper.

Mak, W.K., Morton, D.P. and Wood, R.K. (1999). Monte Carlo Bounding Techniques for Determining Solution Quality in Stochastic Programs. *Operations Research Letters*, 24: 47–56.

Moon, I.K., Ngoc, A.D.D. and Hur, Y.S. (2010). Positioning Empty Containers among Multiple Ports with Leasing and Purchasing Considerations. *OR Spectrum*, 1-22.

Norkin, V.I., Pflug, G.Ch. and Ruszczy´nski, A. (1998). A Branch and Bound Method for Stochastic Global Optimization. *Mathematical Programming*, 83: 425–450.

Steenken, D., Voß, S. and Stahlbock, R. (2004). Container Terminal Operation and Operations Research-a Classification and Literature review. *OR-Spectrum,* 26: 3-49.

Verweij, B., Ahmed, S., Kleywegt, A.J., Nemhauser, G. and Shapiro, A. (2003). The Sample Average Approximation Method Applied to Stochastic Routing Problems: A Computational Study. *Computational Optimization and Applications*, 24(2-3): 289-333.

Zhang, R., Yun, W.Y. and Moon, I. (2009). A Reactive Tabu Search Algorithm for the Multi-depot Container Truck Transportation Problem. *Transportation Research Part E: Logistics and Transportation Review*, 45(6): 904-914.

Zhou, W.H. and Lee, C.Y. (2009). Pricing and Competition in a Transportation Market with Empty Equipment Repositioning. *Transportation Research Part B: Methodological*, 43(6): 677-691.

Session I2: Scheduling 2

· **Day2: Sep. 16 (Thu.)**

· **Time: 16:20 - 17:40**

· **Chair: Byung-In Kim**

· **Room: Azalea & Lilac, 5F**

BERTH AND UNLOADER SCHEDULING PROBLEM AT A RAW-MATERIAL PORT

Byung-In Kim and Sooyoung Chang

Department of Industrial & Management Engineering
Pohang University of Science and Technology, Pohang, Kyungbuk, 790-784, Korea
e-mail: bkim@postech.ac.kr, syc@postech.ac.kr

1. INTRODUCTION

This article addresses the berth and unloader scheduling problem of a steel works of P-company, which is one of the largest steel makers in the world. The steelworks has four major berths with different water depths. Each berth's eligibility for a ship is determined by its water depth and the ship's draft. Thus, when a number of large ships arrive in a short period of time, some of them have to wait and the total duration of each ship's stay at the port becomes longer. When the total duration of a ship's stay is longer than the contracted duration, the company should pay appropriate demurrage cost to the ship owner. In order to reduce the waiting time and the demurrage cost of a ship, the previously berthed ship can be shifted onto another berth after sufficient amount of materials are unloaded and it reaches the water depth of the destination berth. The unloaders have different unloading capacity and eligibility for the berths. Thus, the unloading time of a ship varies depending on which berths and which unloaders are allocated for the ship. The issues of berth allocation, crane scheduling, storage yard allocation, inter-terminal and intra-terminal container transportation at container terminals are reviewed in Vis and Koster (2003) and Steenken *et al.* (2004).

Although the problem considered in this article is specific to the steelworks, it has general and interesting characteristics that are worth to report. We believe that this article extends the previous research on berth allocation and crane scheduling. The ship handling time of our problem depends not only on the assignment of berth but also on the allocation of unloaders and their types. Ships also can be shifted between berths. To the best of our knowledge, no previous research works have considered these characteristics simultaneously. We propose heuristic algorithms for the problem. The performance of the proposed algorithms is evaluated using real data.

2. PROBLEM DESCRIPTION

At the steelworks, ships carrying various raw materials such as coal and iron ore arrive and the materials are unloaded. A typical raw-material ship that arrives at the steelworks is about 300m long with a breadth of 50m, and can carry 200K ton (dead weight tonnage). Each ship has a different draft depending on the amount of materials it contains. The port has four berths: #10, #11, #13, and #14. The water depths of the berths are different. Ships whose draft is less than the water depth of a berth can be berthed on the berth. Roughly speaking, the water depths of berths #10, #11, #13, and #14 correspond to the drafts of ships of 200K ton, 150K ton, 100K ton, and 100K ton, respectively.

There are ten unloading equipments and they have different speed: 6 are on the berths #10 and #11, and 4 are on the berths #13 and #14. There are rails between berths #13 and #14 and between berths #10 and #11 such that the unloaders can move between the berths. However, the maximum number of unloading equipment that can unload materials from a ship at the same time is typically limited to four due to the conflict avoidance criteria among the equipment.

The unloaded raw material by the equipment is then transported to storage yards through a complex belt conveyor network. There are multiple routes from unloading equipment to a storage yard. Since multiple ships can be unloaded and thus multiple routes can be demanded at the same time and some routes might use the same belt conveyor unit, a careful routing scheduling is required in order to keep from delaying due to route conflicts. The route scheduling problem with consideration of the assignment of ship hold and unloading equipment in this environment is a quite complicated and interesting problem. Kim *et al.* (2010) addressed this problem. In this article, we assume that the routing problem can be resolved smoothly following the berthing and unloading schedule. In fact, the planners in the steelworks currently make the berthing schedule only and they dispatch unloading equipment and route in an ad-hoc fashion.

The berth and unloader scheduling problem is to determine the ship's berthing time, unloader assignment, and the unloading completion time of the ship.

3. SOLUTION APPROACH

This section describes our solution approaches for the problem: a heuristic algorithm and an iterative random search algorithm. Both algorithms are very simple and easy to implement. They follow the first-come-first-berthing policy which is the business rule of the company.

3.1. A Heuristic Algorithm

The procedure steps of our first algorithm are as follows:

Step 0. Obtain the desired maximum number of unloaders for each berth from the user.
Step 1. Sort the ships in an ascending order of their arrival times.
Step 2. Do the following steps for each ship.
Step 3. Attempt to berth the ship at each eligible berth, assign appropriate unloaders to the ship and calculate the unloading finish time for the ship at the berth.
Step 4. Find the best berth where the ship can be unloaded at the earliest and make the berthing permanent.
Step 5. If the ship is berthed later than its desired ready time, then do the following steps. Otherwise, go to step 2.
Step 6. Attempt to shift the previous ship that berthed just before the ship at the same berth.
Step 7. Find the best shifting time, the best shifting berth for the previous ship and perform the shifting if it is appropriate.

In our approach, the user can specify his or her desired maximum number of unloaders for each berth (step 0). For example, the user can set 3, 3, 2, and 2 as the maximum numbers for berth #10, #11, #13, and #14, respectively. Then 3, 3, 2, and 2 unloaders are dedicated to berth #10, #11, #13, and #14, respectively. If the user sets 4, 4, 3, and 3 for berth #10, #11, #13, and #14, respectively, then two unloaders can be shared at berths #10 and #11, and another two unloaders can be shared at both #13 and #14. When the user decides that the unloaders are not dedicated to the berths, the unloaders are dynamically assigned to the berths by first-request-first-serviced policy along the procedure. Figure 1 shows an example. CSU1 and CSU2 are assigned to ship #1 at berth #10 first and then they are assigned to ship #2 at berth #11.

The procedure follows the first-come-first-berthing policy and thus the ships are sorted in an ascending order of their arrival times (step 1) and are berthed in that order (step 2). In step 3, each ship is tried to be berthed at each eligible berth, while appropriate unloaders are assigned to the ship, and the finish time for the ship is calculated. For example, Figure 2(a) shows that three ships have already berthed and ship #4 is arriving at the port (notice that the unloaders are dedicated onto the berths in this example). Suppose that all the berths are eligible to the ship. Then the ship is tried to be berthed at the berths and the finish time for each case is calculated as shown in Figure 2(b). When the finish time of a ship is estimated at each berth, the maintenance schedule and existing schedule, if there are any, are taken into consideration. Then the best berth in which the ship can be unloaded as soon as possible is selected (step 4) and the berthing is made permanent. In this example, berth #10 would be selected since it provides the earliest completion time for the ship.

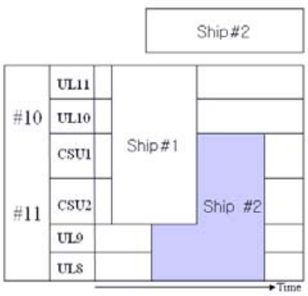


Figure 1. Dynamic unloader assignment

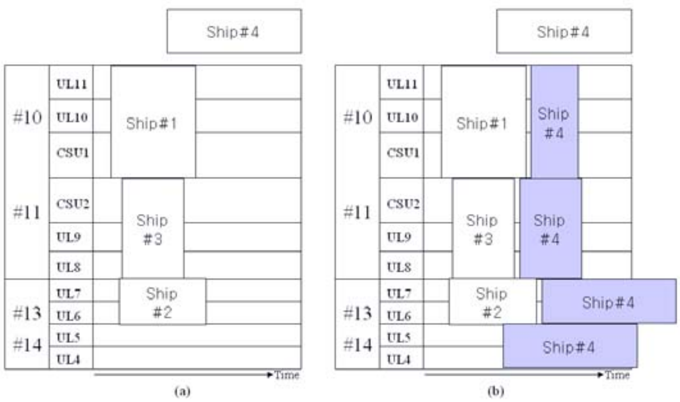


Figure 2. Berthing example

In step 5, the ship's desired ready time and its actual berthing time are checked. If the ship is berthed later than its desired time, then shifting of the previous ship at the berth is considered (step 6). Figure 3 shows an example. In Figure 3(a), we can see that ship #4 is berthed later than its ready time due to ship #1. Then, the algorithm attempts to shift ship #1 onto other berths. Figure 3(b) shows the result when ship #1 is shifted onto berth #11 and Figure 3(c) shows one when it is shifted onto berth #14. The shifting of the ship onto berth #13 does not need to be considered because there is no benefit to do so in this case. In step 7, the best shifting time and the best shifting berth are selected and the shifting is made permanent. In the example case, we choose Figure 3(c) as the best because the shifting makes ship #4 berth at the earliest. Of course, we can use different criteria to choose the best shift considering other effects. When a ship is to be berthed or is to be shifted, possible piloting time should be considered. Normally, piloting is not allowed during the night time at the port.

Notice that a ship can be shifted entirely from a berth to another berth with this mechanism. Figure 4 shows an example. In this case, ship #3 could have been berthed at berth #14 from the beginning but since the completion time of the ship at berth #11 was earlier than at berth #14, it was berthed at #11. Now, ship #4 is to be berthed and it is eligible to berth only at berth #10 and #11 due to the draft and water depth constraints. Once it is berthed at berth #11, ship #3 is considered to shift. Since ship #3 can be shifted entirely and it is desirable to do so, the ship is berthed at berth #14.

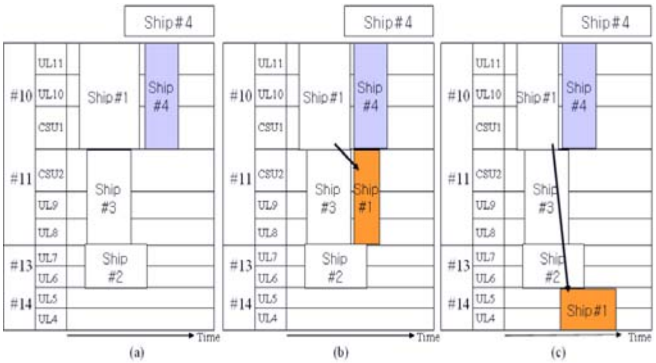


Figure 3. Shifting example 1

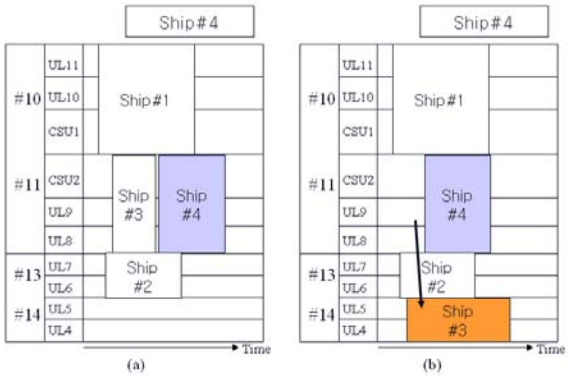


Figure 4. Shifting example 2

3.2. An Iterative Random Search Algorithm

The heuristic algorithm described in the previous subsection is a type of sequential greedy heuristics. In most sequential greedy heuristics, early decisions affect all the subsequent decisions. Although a decision is considered as the best with a myopic view, it would turn out to be a poor decision later on. In step 4 of the previous algorithm, a berth in which the ship can be unloaded the earliest was selected to be the best without considering the future events. In this subsection, we modify the algorithm and present an iterative random search algorithm in which an eligible berth is randomly selected and the procedure is repeated. The procedure steps of the modified algorithm are as follows:

Step 0. Obtain the maximum iteration number and maximum computation time from the user.
Obtain the desired maximum number of unloaders for each berth from the user.
Initialize the best solution with the heuristic approach of section 3.1.
Step 1. Sort the ships in ascending order of their arrival times.
Step 2. **Do the following steps until the maximum iteration number or maximum computation is reached.**
Step 3. Do the following steps for each ship and make a new solution.
Step 4. **Select an eligible berth randomly and make the ship berth.**
Step 5. If the ship is berthed later than its desired ready time, then do the following steps.
Otherwise, go to step 3.
Step 6. Attempt to shift the previous ship that berthed just before the ship at the same berth.
Step 7. Find the best shifting time, the best shifting berth for the previous ship and perform the shifting if it is appropriate.
Step 8. **If the new solution is better than the current best solution, then update the best solution with it.**

The highlighted steps 2, 4, and 8 are the different steps from the previous heuristic algorithm. The algorithms are compared in the next section. In step 4, a berth is selected randomly among the eligible berths. In step 8, the best solution is updated whenever a better solution is obtained.

4. PERFORMANCE EVALUATION

The algorithms were implemented in Microsoft Visual C++ 6.0 and a prototype system has been developed and tested on several real-world problems. For a one-month data set, while the actual activities performed by the company caused $115,699 of demurrage cost and $96,699 of earlier discharging revenue, our first heuristic generated a solution whose demurrage cost was $166,943 and earlier discharging revenue was $156,608. The result of the iterative random search algorithm with the maximum number of iterations of 10,000 caused the demurrage cost of $76,522 and earlier discharging revenue of $168,972. The iterative random search algorithm outperformed the first heuristic algorithm. The computation times for the first algorithm and the iterative one were less than 1 second and about 80 seconds, respectively, on a Windows XP Pentium 4, 3.4 GHz processor with 3.25 GB of RAM.

5. CONCLUSIONS

This article described a complicated berth and unloader scheduling problem in which berths have different water depth, and the ship handling time depends not only on the assignment of berth but also on the allocation of unloaders and their types. Furthermore, the ships can be shifted between berths, and the shifting moment is not given but is a decision to make. We proposed two heuristic algorithms. The proposed algorithms were implemented in a prototype system and the system has been tested using the actual data. The proposed algorithms outperformed the actual results. More detailed description of the problem and results are in Kim and Chang (2009) and Kim *et al.* (2010).

6. REFERENCES

Kim, B. and Chang, S. (2009). Raw-material unloading at a steelworks. *Proceedings of 2009 ICOVACS (International Conference on Value Chain Sustainability)*, Oct 19-21, 2009, Louisville, KY, USA

Kim, B., Chang, S., Chang, J., Han, Y., Koo, J., Lim, K., Shin, J., Jeong, S. and Kwak, W. (2010). Scheduling of raw-material unloading from ships at a steelworks. *Production Planning & Control*, To appear.

Steenken, D., Vob, S. and Stahlbock, R. (2004). Container terminal operation and operations research – a classification and literature review. *OR Spectrum*, 26 (1), 3-49.

Vis, I.F.A. and de Koster, R. (2003). Transshipment of container at a container terminal: an overview. *European Journal of Operational Research*, 147, 1-16.

AN EXACT ALGORITHM FOR THE BERTH ALLOCATION PROBLEM WITH TIME WINDOWS

K.L. Mak[*] and Sun Di

Department of Industrial and Manufacturing Systems Engineering
The University of Hong Kong
Pokfulam Road
Hong Kong
e-mail: makkl@hkucc.hku.hk,

Abstract: This paper studies the problem of assigning discrete berths to incoming vessels with time windows. The lower bound and the upper bound of the time window is the vessel arrival time and the vessel completion due time, respectively. Vessels could only moor at certain berths due to physical or technical restrictions, such as the relationships between ship draft and water depth, long-term contracts or specific handling equipment. The ship handling time at the berth is not only dependent on the size of the ship, but also the location where it moors. The objective is to minimize the total port stay times of ships. This paper proposes an exact algorithm called Hybrid Branch and Price with Heuristics (HBPH) algorithm to locate the global optimal solutions. Computational experiments show that HBPH can locate the global optimal solutions for real-size problems within acceptable times.

1. INTRODUCTION

In a container terminal, berths are the most critical resources for determining the terminal capacity, because the cost of constructing a berth is very high compared to the investment costs of the other facilities in the terminal. Hence, efficient utilization of berths is essential to improve the productivity of the container terminal.

The method of allocating berth space for ships in a container terminal is called the berth allocation problem (BAP). To ensure maximum container terminal throughput, terminal planners often have to make two interrelated decisions of where and when the ships should moor. The problem can be represented in a two-dimensional decision space where a berth-allocation is represented by a rectangle. Figure 1 illustrates the allocation of a ship to a particular berthing position for a specified handling time. The dimensions are length (with a safety margin) and handling times. Hence, the rectangles representing the allocations of ships to berths along the quay must be placed in the decision space without overlapping each other while at the same time, satisfying the various operational constraints.

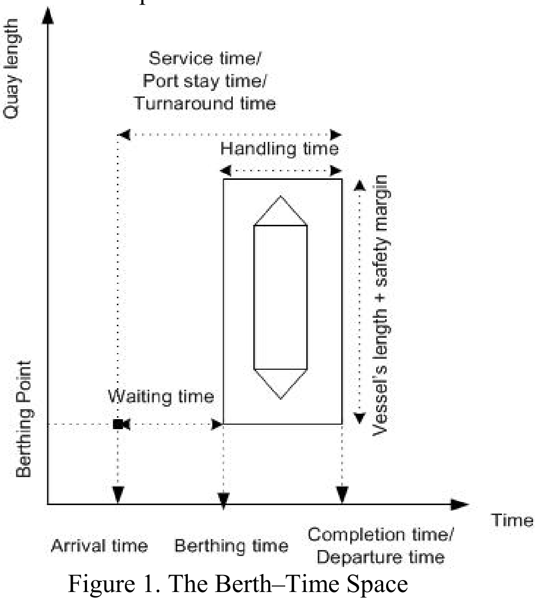


Figure 1. The Berth–Time Space

Berth allocation problems can be classified as static or dynamic. The static problem (SBAP) treats ships that have already arrived at the port and wait for berthing (Imai *et al.*, 1997). These ships could be berthed as long as the berths are available. This is a fairly unrealistic assumption. To avoid this shortcoming, Imai *et al.* (2001) introduced the dynamic version of the berth allocation problem. Their paper not only takes into account the ships that have already arrived before planning but also ships which will arrive in some specified future time.

Berth allocation problems are often modeled in a discrete form according to the berth structure. Allocation of ships to discrete berthing positions in a container terminal (i.e., a berth allocation problem, BAPD) considers the quay as a finite set of berths. Under this scenario, the berths can be described as fixed length segments, each of which could moor one ship at a time and each ship should be assigned to one berth only. In particular, there may be time windows which govern the vessel arrival times and the vessel completion due times. In this paper, a mathematical model which represents the characteristics of the problem, and an efficient algorithm are developed to minimize the total port stay times of ships.

2. BAPTW MODELING

The problem considers the allocation of ships with different time windows to discrete berths subject to various operational restrictions, in order to minimize the total port time of ships. The following assumptions are made in the development of the mathematical model:

- Every ship must be serviced exactly once at any berth.
- Each berth can only serve one ship at a time.
- Some ships could only moor at several certain berths due to physical or technical restrictions, such as the relationships between ship draft and water depth, long-term contracts or specific handling equipment. If a ship could not moor at a berth, the ship handling time at the berth is set to positive infinity.
- Once a ship could moor at a berth, the ship handling time at the berth is dependent on not only the size of the ship but also the location where it moors.

The model can be represented by a multi-graph $G^k = (V^k, A^k), \ \forall k \in M$ where $V^k = N \cup \{o(k), d(k)\}$ and $A^k \in V^k \times V^k$. The notations used in the mathematical model are defined as follows:

N : The set of ships, $n = |N|$;

M : The set of berths, $m = |M|$;

h_i^k : Handling time of ship i at berth k;

a_i : Arrival time of ship i;

d_i : Department due time of ship i;

l_k : Lower bound of available time window of berth k;

u_k : Upper bound of available time window of berth k;

$$x_{ij}^k = \begin{cases} 1 & \text{if the ship } j \text{ is scheduled after ship } i \text{ at berth } k \\ 0 & \text{otherwise} \end{cases};$$

T_i^k : The berthing time of ship i at berth k.

Hence, the characteristics of the berth allocation problem can be represented by the BAP model:

$$\min \sum_{i \in N} \sum_{k \in M} (T_i^k - a_i + h_i^k) \sum_{j \in N \cup \{d(k)\}} x_{ij}^k \qquad (1)$$

Subject to:

$$\sum_{k \in M} \sum_{j \in N \cup \{d(k)\}} x_{ij}^k = 1 \qquad \forall i \in N \qquad (2)$$

$$\sum_{j \in N \cup \{d(k)\}} x_{o(k)j}^k = 1 \qquad \forall k \in M \qquad (3)$$

$$\sum_{i \in N \cup \{o(k)\}} x_{id(k)}^{k} = 1 \qquad \forall k \in M \tag{4}$$

$$\sum_{j \in N \cup \{d(k)\}} x_{ij}^{k} - \sum_{j \in N \cup \{o(k)\}} x_{ji}^{k} = 0 \qquad \forall k \in M \ \forall i \in N \tag{5}$$

$$T_i^k + h_i^k - T_j^k \le (1 - x_{ij}^k)M \qquad \forall k \in M, \ \forall i, j \in N \tag{6}$$

$$T_i^k \ge a_i \qquad \forall i \in N \tag{7}$$

$$T_i^k + h_i^k \le d_i \qquad \forall i \in N \tag{8}$$

$$T_{o(k)} \ge l_k \qquad \forall k \in M \tag{9}$$

$$T_{d(k)} \le u_k \qquad \forall k \in M \tag{10}$$

The objective function minimizes the total port stay time of ships. Constraints (2) state that each ship is assigned to one berth only once. Constraints (3) and (4) ensure each berth will handle the first and last ship. The flow conservation is given in constraints (5). Constraints (6) calculate the ships berthing times. Constraints (7) state berthing times of ships must be posterior to the ships arrival times. Constraints (8) assure the ships must leave the port before their completion due times. Constraints (9) and (10) are the time windows for each berth.

The Berth Allocation Problem described above could be seen as a specific case of Multiple Depot Vehicle Routing Problem with Time Windows (MDVRPTW), where the number of vehicles is fixed. The ships and berths are corresponding to customers and depots, respectively. There are m vehicles, one for each depot, and each vehicle starts and finishes its tour at its depot. Kolen *et al.* (1987) has shown that even finding a feasible solution to the Vehicle Routing Problem with Time Windows (VRPTW) when the number of vehicles is fixed is a NP-complete problem. Thus, it is difficulty to determine whether a given berth allocation problem described above has feasible solutions. In practice, if the berth capacities are so limited that cannot satisfy all the time window constraints, port managers have to delay departure time of some vessels and provide compensation to the companies which own those vessels. Thus, this paper relaxes the time constraints (8) and (10) and adds them to the objective function with weight factors (ω_1 and ω_2) for punishment of violating time constraints (8) and (10). The objective function in this case is modified as

$$\min \ (\sum_{i \in N} \sum_{k \in M} T_i^k - a_i + h_i^k + \omega_1 \cdot \max(0, T_i^k + h_i^k - d_i)) \sum_{j \in N \cup \{d(k)\}} x_{ij}^k + \omega_2 \cdot \max(0, T_{d(k)}^k - u_k) \ .$$

3. AN EXACT ALGORITHM BASED ON COLUMN GENERATION METHODS

There is no exact algorithm in the literature for solving discrete berth allocation problem. The research objective of this study is therefore to design effective and efficient exact algorithms based on column generation approaches to identify the global optimal solution to BAPD. To introduce the algorithm in details, this paper first reformulates the discrete berth allocation problem by using a partition model with columns. Each column is represented by a serial of integers, where the first position indicates the berth referring to the column, and the other positions represent the ship handling sequence of the berth (column). The notations are defined as follows:

$$b_{ik} = \begin{cases} 1 & \text{if the column } k \text{ attends the ship } i \\ 0 & \text{otherwise} \end{cases};$$

$$e_{jk} = \begin{cases} 1 & \text{if the column } k \text{ represent berth } j \\ 0 & \text{otherwise} \end{cases};$$

P : Set of all feasible columns, $p = |P|$;

c_k : The cost of column k.

Master problem: MP

$$\min \ \sum_{k=1}^{p} c_k x_k \tag{11}$$

Subject to:

$$\sum_{k=1}^{p} b_{ik} x_k = 1 \qquad \forall i \in N \tag{12}$$

$$\sum_{k=1}^{p} e_{jk} x_k \le 1 \quad \forall j \in M \tag{13}$$

$$x_k = \begin{cases} 1 \text{ if column } k \text{ is used} \\ 0 \ \ \text{otherwise} \end{cases} \qquad \forall k \in P \tag{14}$$

Where $c_k = \sum_{i \in B_k} (T_i^k - a_i + h_i^k) + \sum_{i \in B_k} \omega_1 \cdot \max(0, T_i^k + h_i^k - d_i) + \omega_2 \cdot \max(0, T_{d(k)}^k - u_k)$, B_k is the set of ships attended by column k. Objective (11) is to minimize the total ship stay times. Constraints (12) ensure all ships are handled by one berth only once. Constraint (13) states the columns used in a feasible solution must represent different berths. To relax the integer constraints of (14), we get the linear relaxation master problem (LMP).

Since the set of all feasible ship sequences is extremely large, one cannot expect to generate it completely. To solve the linear relaxation master problem without enumerating all the ship sequences, a column generation technique is proposed. Specifically, this is done by enumerating a partial set of ship sequences P' and $p' = |P'|$. The formulation of the corresponding restricted linear relaxation master problem with respect to this set is as follows:

Restricted linear master problem: RLMP

$$\min \sum_{k=1}^{p'} c_k x_k$$

Subject to:

$$\sum_{k=1}^{p'} b_{ik} x_k = 1 \qquad \forall i \in N$$

$$\sum_{k=1}^{p'} e_{jk} x_k \le 1 \quad \forall j \in M$$

$$x_k \ge 0 \qquad \forall k \in P'$$

Let $\bar{x}_k \ \forall k \in P'$ be the optimal solution to RLMP, and let $\bar{\pi}_i \ \forall i \in N$ and $\bar{\mu}$ be the corresponding optimal dual variables. We would like to know whether $\bar{x}_k \ \forall k \in P'$ is optimal for the linear relaxation master problem (LMP), and equivalently, $\bar{\pi}_i \ \forall i \in N$ and $\bar{\mu}$ is the dual solution of the LMP. To answer this question, we should observe the dual formulation of the LMP.

Dual problem of the linear relaxation master problem: MPD

$$\max \sum_{i=1}^{n} \pi_i + \sum_{j=1}^{m} \mu_j$$

Subject to:

$$\sum_{i=1}^{n} b_{ik} \pi_i + \sum_{j=1}^{m} e_{jk} \mu_j \le c_k \quad \forall k \in P$$

$$\mu_j \le 0 \quad \forall j \in M$$

Clearly, if $\overline{\pi}_i \ \forall i \in N$ and $\overline{\mu}$ satisfies every constraint in MPD, they are dual optimal for the linear relaxation master problem and therefore the solution $\overline{x}_k \ \forall k \in P'$ is proved optimal for the LMP. The pricing cost of column k denoted as r_k is defined as follows:

$$r_k = c_k - \sum_{i=1}^{n} b_{ik}\pi_i - \sum_{j=1}^{m} e_{jk}\mu_j$$

$$\because b_{ik} = \sum_{j \in N \cup \{d(k)\}} x_{ij}^k$$

$$\therefore r_k = \sum_{i \in N} (T_i^k - a_i + h_i^k + \omega_1 \cdot \max(0, T_i^k + h_i^k - d_i) - \pi_i) \sum_{j \in N \cup (d(k))} x_{ij}^k + \omega_2 \cdot \max(0, T_{d(k)}^k - u_k) - \sum_{j \in M} e_{jk}\mu_j$$

$$= \sum_{i \in B_k} (T_i^k - a_i + h_i^k + \omega_1 \cdot \max(0, T_i^k + h_i^k - d_i) - \pi_i) + \omega_2 \cdot \max(0, T_{d(k)}^k - u_k) - \sum_{j \in M} e_{jk}\mu_j$$

If the columns with price $r_k < 0$ could be detected, their corresponding constraints in dual problem would be violated. In that case the current solution $\overline{\pi}_i \ \forall i \in N$ and $\overline{\mu}$ is not optimal for the problem MPD. The columns just found can be added to the formulation of problem RLMP, which is solved again. The process repeats itself until no violated constraint is found. In this case this procedure locates the optimal solution of the linear relaxation master problem which is also the optimal solution of problem MPD. Now the task focuses on finding columns with negative pricing cost. Each berth has a corresponding column generation sub-problem due to specific features of the berth.

Column generation sub-problems:

$$z^k = \min \left(\sum_{i \in B_k} T_i^k - a_i + h_i^k + \omega_1 \cdot \max(0, T_i^k + h_i^k - d_i) - \pi_i\right) + \omega_2 \cdot \max(0, T_{d(k)}^k - u_k) - \sum_{j \in M} e_{jk}\mu_j$$

Subject to:

$$\sum_{j \in N \cup \{d(k)\}} x_{o(k)j}^k = 1$$

$$\sum_{i \in N \cup \{o(k)\}} x_{id(k)}^k = 1$$

$$\sum_{j \in N \cup \{d(k)\}} x_{ij}^k - \sum_{j \in N \cup \{o(k)\}} x_{ji}^k = 0 \qquad \forall i \in N$$

$$T_i^k + h_i^k - T_j^k \le (1 - x_{ij}^k)M \qquad \forall i, j \in N$$

$$T_i^k \ge a_i \qquad \forall i \in N$$

$$T_{o(k)} \ge l_k \qquad \forall k \in M$$

Since the structures of berths are different from each other. There are m column generation sub-problems each of which corresponds to a berth. Each sub-problem only generates the columns for the corresponding berth. The objective of each column generation subproblem is to find columns with negative price cost ($z^k < 0$). This subproblem could be seen as the elementary shortest path problem with time windows constraint (ESPPTW) in graphs that might contain negative cost paths. The exact solution method of ESPPTW is introduced in the next subsection.

3.1 Elementary Shortest Path Problem with Time Windows (ESPPTW)

This paper designs an exact approach based on label-correcting method to find the optimal solutions of the column generation subproblem. Each ship is denoted as a node. The berth sequence for each berth could be considered as a traveling path. Thus, the ship sequence (column) generation problem is an elementary shortest path problem with time windows. Node i corresponding to the ship served by berth k has labels: $l_i^k = (k, T_i, C_i, S_i, V_1, \ldots V_n)$. Label l_i^k indicates a path from the origin node $d(k)$ of berth k to node i, where ($V_1, \ldots, V_n$) corresponds to the node sequence travelled in the path. For example, there are 4 nodes and the node indicating vector is $(V_1, V_2, V_3, V_4) = (1, 0, 2, 3)$, which means the path is (Node 1, Node 3, Node 4). Based on a partial path, some nodes called unreachable nodes could be eliminated from future visit due to time constraints or other physical restrictions. In this case, the corresponding elements V_j in node indicating vector are set to $n+1$.

P_i: The set of node j where $V_j = 0$ in label l_i^k;

L_i: The set of non-dominated labels of node i;

W: List of labels waiting to be treated;

C_i: The path cost from origin node $o(k)$ to node i;

S_i: The number of ships travelled by label l_i^k;

T_i: The time when the path arrives node i.

The definition of dominance rule is that L_i^1 dominates L_i^2 if and only if $k^1 = k^2$, $C_i^1 < C_i^2$, $S_i^1 < S_i^2$.

The algorithm procedure, which determines all the un-dominated paths starting from node $o(k)$ to every node of the graph for berth k is described as follows:

Initialize

$l_{o(k)}^k : (k; T_{o(k)} = 0; C_{o(k)} = 0; \forall j \in N$, if ship j could be handled at berth $k, V_j = 0$; else, $V_j = n+1)$

$W = \{l_{o(k)}^k\}$

Treat $l_{o(k)}^k$

For all successors $j \in P_{o(k)}$

{

Extend $l_{o(k)}^k$ to get l_j^k:

$T_j = \max(l_k, a_j)$, $C_j = h_j^k - \pi_j$;

Set $V_j = 1$, $S_j =$ the number of V_j, where $V_j \neq 0$;

Set $V_q = n+1$, if q is an unreachable node;

Get the successors set P_j of label l_j^k;

Set $L_j = \{l_j^k\}$, $W = W \cup \{l_j^k\}$.

}

Set $W = W \setminus \{l_{o(k)}\}$

Repeat

Treat $l_i^k \in W$

For all successors $j \in P_i$:

{

Extend l_i^k to get new label $l_j^{k\prime}$ for node j:

$T_j' = \max\{a_j, T_i + h_i^k\}$;

$C_j' = C_i + T_j' - a_j + h_j^k - \pi_j$;

Set $V_j' = V_i + 1$, V_q' $\forall\{q \in N \quad j\}$ is the same as in l_i^k;

Set $V_q' = n+1$, if q is an unreachable node;

Get the successors set P_j' of label $l_j^{k\prime}$;

Set $S_i' =$ the total number of V_q', where $V_q' \neq 0$ $\forall q \in N$.

For all labels $l_j^k \in N_j$

{

If label $l_j^{k'}$ is an un-dominated label of

node j, set $L_j = L_j \cup \{l_j^{k'}\}$, $W = W \cup \{l_j^{k'}\}$

In addition, if label $l_j^{k'}$ dominates label l_j^k,

set $L_j = L_j \setminus \{l_j^k\}$

}

}

Set $W = W \setminus \{l_i^k\}$

Until $W = \phi$

3.2 Hybrid Branch and Price Algorithm with Heuristics (HBPH):

Two heuristics, the assignment heuristic and construction heuristic (A-C heuristic) are employed to generate good integer solutions early. Assignment heuristic distributes each ship to its favorite or acceptable berths. It creates initially m empty berths. The n ships are sorted by arrival order. Each ship i is distributed to a berth $k \in FAB_i$ in a random way, where FAB_i is the corresponding berth set of ship i including its favorite and acceptable berths. Construction heuristic makes the ships schedule in the berths. The berthing time of each ship and the objective function value of each column are computed in this heuristic.

There are two different purposes to apply the A-C heuristic: the first one is to generate initial columns set to form a feasible solution for model RLMP. The second one is to combine A-C heuristic with the above label-correction algorithm to solve the column generation subproblem. A-C heuristics are firstly used to generate columns until no columns with negative pricing cost are found. Then, the program would call for label correcting algorithm to continuously solve the column generation subproblem.

3.3 Branching Schemes

An optimal solution of the LP relaxation may contain variables with fractional values. Applying a standard branch and bound procedure does not guarantee optimality because the column set does not contain all of the variables. Therefore a branch and bound framework has to be established. The calculations are organized in a branching tree. The branching decisions are generally based on considerations related to the original BAP model. The column generation process is then repeated at each node in the branch and bound tree.

Branching can be done on sums of flows $\sum_{k \in M} x_{ij}^k$. The branching decision can easily be transferred to the master problem and sub-problems. When $\sum_{k \in M} x_{ij}^k = 1$, customer j succeeds customer i on the same route, while if $\sum_{k \in M} x_{ij}^k = 0$, customer i does not immediately precede j. Each arc (i, j) of columns is evaluated by a score which is calculated as the sum of flow values on that arc. The arc with the best score is chosen on branching.

4. COMPUTATION EXPERIMENT

4.1 Test Problem Generation:

Problems used in these experiments are generated randomly, but systematically. Different parameter settings of berths and ships yield various prototype problems. The favorite berth of a ship i, denoted as i_k is randomly picked up from the berth set M ($i_k \in M$). Besides i_k, this paper further randomly selects $\alpha \in \{0,1,2,3\ldots m\}$ berths called acceptable berths, which could also service the ship i. In this study, the total number of acceptable berths and favorite berth is about the 80% of total number of berths. At first, the handling time of a ship at its favorite berth $h_i^{i_k}$ is generated from a 2-Erlangian distribution. Then, the handling time of the ship at an acceptable berth is determined by the $h_i^{i_k}$ and the distance between the acceptable berth location and the favorite berth location. In other words, the longer distance between the acceptable berth and the favorite berth, the longer handling time of the ship at the acceptable berth. Thus, the handling time of ship i at an acceptable berth k is calculated as $h_i^k = h_i^{i_k} + \beta |k - i_k| \cdot h_i^{i_k}$ where $0 < \beta < 1$. The handling time of the ship at other unacceptable berths is set to positive infinity.

The interval of ship arrivals obeys the exponential distribution (Imai *et al.* (2001)). The completion due time of a ship is set as: $d_i = a_i + \lambda \cdot h_i^{i_i}$ where $\lambda > 1$. The earliest available time of berth k is obtained as in Imai *et al.* (2001) and Cordeau *et al.* (2005): l_k is the same for every berth; l_k is set for each instance equal to a given fraction f of the time duration between the first and last arriving ships. The parameter f represents the level of an instance close to the "static" version of BAP. When $f = 1$, the dynamic BAP is reduced to static BAP. In this study, f is fixed as $\frac{1}{12}$. The latest available time of berth k is the same for every berth and set as $u_k = d_n$ where n represents the last arrival ship.

4.2 Computational Performances

Table 1 shows the results of solving the problem by using HBPH. Heu-Cols give the number of columns constructed by the heuristics in each sub problem. AvgNodes and Avg-SPs are the average numbers of solved nodes and sub-problems, respectively. Avg-Cols show the total number of columns with negative pricing costs in the whole computation process and Avg-Time is the average computational times. Each test instance is solved three times and the average of the results is used to eliminate the randomness. The results clearly show that the exact algorithm could solve the real problems optimally within acceptable computational times.

Table 1. Solutions found by HBPH

Ins V	essel	Berth	*Acc*Berth	*Heu*Cols O	PT	*Avg*Nodes	*Avg*Cols	*Avg*SPs A	*vg*Time
1 20		5	4	500	615.119	1	2265	9	13.002
2 20		5	4	500	722.716	1	1639	11	19.143
3	20	5	4	500	692.422 1		2190	16 18	.328
4	25	8	6	500	795.413 1		3155	15 30	.179
5 25		8	6	500	1127.78	1	3244	18	36.126
6 30		8	6	600	1101.38	1	3719	20	58.554
7	30	8	6	600	1317.54 1		3562	27 88	.275
8	30 10		8 6	00	1136.89	1	3863	17	55.868
9	30 10		8 6	00	1124.55	1	3815	19	61.845
10	40 10		8 8	00	1892.24	10	30004	306	275.223
11	40 10		8 8	00	1499.68	13	30351	370	231.468
12	40	13	10	800 14	38.85	7	40350	105 19	9.723
13	40	13	10	800 16	51.86	9	40149	125 19	1.272
14 50		13	10	1200	2117.21	21	90675	369	848.435
15 50		13	10	1200	1824.23	23	90740	454	875.407
16 50		15	12	1200	2011.32	15	30064	401	900.825
17 50		15	12	1200	1699.72	17	20494	218	463.129

5. REFERENCES

Cordeau, J.F., Laporte, G., Legato, P., and Moccia, L. (2005). Models and tabu search heuristics for the berth-allocation problem, *Transportation Science*, 39: 526–538.

Imai, A., Nagaiwa K. and Chan, W.T. (1997). Efficient planning of berth allocation for container terminals in Asia. *Journal of Advanced Transportation*, 31: 75-94.

Imai, A., Nishimura, E. and Papadimitriou, S. (2001). The dynamic berth allocation problem for a container port, *Transportation Research Part B*, 35(4): 401–417.

Kolen, A. W. J., Rinnooy Kan, A. H. G. and Trenekens, H. W. J. M. (1987). Vehicle routing with time windows, *Operations Research*, 35(2): 266-273.

OPTIMISING CONTAINER STORAGE PROCESSES AT MULTIMODAL TERMINALS

Brad Casey and Erhan Kozan*

Mathematical Sciences
Queensland University of Technology
GPO Box 2434, QLD 4001, Australia
Email: bv.casey@student.qut.edu.au
e.kozan@qut.edu.au

Abstract: Multimodal Container Terminals (MMCT) requires comprehensive planning, programming and control in order to operate an efficient storage system. In this paper, two models have been developed for the storage system of the MMCT. The first model focuses on a single storage area, where containers are stored for a time while waiting to be moved to other areas in the MMCT. It is an extension of the Blocks Relocation Problem, with incoming as well as outgoing containers, or "blocks". The model deals with assigning containers to positions within the storage area, rehandling of containers to retrieve and calculation of the processing times in order to perform these actions. The second model focuses on the movement of containers between the various storage locations within the MMCT, such as the berth, the train and truck loading stations and the general storage bays within the terminal. It deals with the scheduling of numerous machines in parallel, as well as path finding and collision detection/avoidance for these machines. Algorithms have been designed to solve these modes.

CONTAINER STORAGE PLANNING BY META-HEURISTICS

Etsuko Nishimura[1*], Akio Imai [1], and Yoshiyasu Urushibara[1]

[1] Graduate School of Maritime Sciences
Kobe University
5-1-1 Fukae-minami, Higashinada
Kobe 658-0022, JAPAN
E-mail: e-nisi@maritime.kobe-u.ac.jp

Abstract: This study addresses the container storage arrangement on the container yard for the transshipment containers, in order to carry out the ship handling operations efficiently. We consider containers flow into a terminal in one way from the mega-containership to the feeders, in order to obtain the solution quality in this study. To facilitate the solution procedure we employ the heuristics based by subgradient method with lagrangian relaxation (LR), genetic algorithm (GA) and tabu search (TS). As the results, the earlier the mega-containership arrives after most of feeder ships arrive, the better the solutions obtained by LR are. If the mega-containership arrives after a half of feeder ships arrive, TS is better than others in the linear type, GA is better than others in the indented type. Best algorithm depends on the arrival time of the mega- containership or the terminal layout.

1. INTRODUCTION

There has been shown to be a tendency towards great numbers of the container throughput as the world trade expands. However, since Lehman Shock has occurred on September 2008, the major ports in Asia also have been seriously affected. The container throughput of 2009 is less than that of 2008, decreasing rates in Hong Kong (from Jan. to Sep.) and Singapore (from Jan. to Dec.) are 17.3% and 13.5%, respectively. After that, since beginning of 2010 there is a tendency towards great numbers of the container throughput, that trend becomes recovering. The increasing rates for container throughput between March 2009 and March 2010 in Hong Kong and Singapore are 11% and 9.9 %, respectively. Hereafter it is expected to change a trend as stated above, mega-containerships with the capacity of over 10,000 TEU are serviced between Asia and Europe. There are around twenty those ships as of 2008. From the website of Maersk, looking at some ships belonging to shipping company "Maersk" out of those vessels, those ships do not stay at the ports so long time period. We can estimate that the number of containers for loading and discharging is not so much. The number of ports where mega-containerships call has been almost same as conventional size ships yet. However, in the future, the effect on introducing mega-containerships will become apparent. At a terminal where mega-containerships are calling, most cargoes to be handled are transshipments from origin to destination ports via that terminal. Thus, associated handling operations are undertaken between mega-containerships and feeder ships. In order to decrease the turnaround time of ships calling at this terminal, smooth handling of operations is of major importance.

This study is concerned with the container storage planning problem at the terminal, in order to carry out the container handling operations in the most efficient way. This paper is organized as follows. The next section reviews the related literature. In Section 3 the problem is defined. In sections 4 and 5, the formulation for the problem and the solution procedure are described, respectively. In section 6 numerical experiments are carried out and results are presented. The final section reports the paper's findings and conclusions.

2. LITERATURE REVIEW

Storage space is a critical resource in container terminals. Kozan and Preston (1999) consider the determination of optimal storage strategies and container handling schedules by yard cranes. They propose a heuristic method by a genetic algorithm (GA) and analyze the factors that influence container transfer efficiency through lower throughput time at a container terminal with different types of handling equipment, storage capacities and layouts.

Kim *et al.* (2003) propose a solution method for determining the storage location of export containers considering their weights. They consider the stack arrangement and the weight distribution of containers in the yard bay. They develop a dynamic programming model for determining the storage allocation, in order to minimize the number of relocation movements. Kim and Park (2003) discuss how to allocate the storage space for outbound containers. They formulate the model as a mixed-integer linear program in order to allocate the space for utilizing space efficiently and make loading operations more efficient. Two heuristic algorithms are proposed to solve this problem.

Kim and Hong (2006) address the problem of relocating containers in yard block. They proposed a branch and bound algorithm and a heuristic rule based on probability theory. And they also proposed a procedure for estimating the expected additional number of relocations which can apply various configurations of stacks. Yang and Kim (2006) address the storage space to containers assignment problem so that the total expected number of relocations is minimized. Container treatment is grouped as storage demand unit by destination port, weight and size classes. They consider the static problem in which arrival times and departure times of containers are known, the dynamic problem in which arrival times of containers are not known. They solve the static problem using a dynamic programming technique and a genetic algorithm, the dynamic problem by their proposed heuristic methods.

Zhang *et al.* (2003) address the storage space allocation problem in the yard. They solve the problem using a rolling-horizon approach. This problem is decomposed into two levels: placing the total number of containers in each block at the time period, and determining the number of containers associated with each vessel in order to minimize the total distance to transport the containers between yard blocks and the ship berthing location. They deal with four types of containers as follows: inbound containers on vessels and already stored in yard blocks, outbound containers grounding in yard blocks and already stored in yard blocks. However Bazzai *et al.* (2009) consider only inbound containers at the first level of the problem that Zhang *et al.* (2003) dealt with, they determine the number of inbound containers of each vessel stored in each block, in order to minimize the imbalance among workloads allocated to blocks.

Lee *et al.* (2006) consider the problem in order to determine the minimum number of yard cranes to deploy and the location where the group consisted by unloaded containers should be stored. They formulated the problem as a mixed integer-programming model, in order to minimize the total number of shifts required to handle all the workload. They also proposed two heuristics: a sequential method and a column generation method. Han *et al.* (2008) consider the container-to-yard location assignment problem, as given the concept to reserve dedicated clusters for each vessel. They determine the number of containers discharged during each period, in order to minimize the total number of yard cranes required to handle the workloads. Lim *et al.* (2006) propose a meta-heuristic procedure, named the critical-shaking neighborhood search, in order to minimize the yard space needed, as given for yard spaces to requests at each time.

Park and Seo (2009) consider the planar storage location assignment problem which can be defined as the assignment of the inbound and outbound containers to the storage yard with aim of minimizing the number of obstructive container moves. They proposed an efficient genetic algorithm to solve the problem for real-sized instances. Park and Seo (2010) also consider the problem as well as Park and Seo (2009). They proposed a genetic algorithm based heuristic and their original heuristic methods. In order to consider three conditions that make blocks move as following: (a) retrieving due outbound blocks from the yard, (b) replacing obstructive blocks in the path of retrieving the due out bound block to other empty space, and (c) storing inbound blocks at the yard, their original heuristic is consisted by two independent sub-algorithms: (1) obstructive block selection and (2) storage location determination.

Summarizing the related works, the container treatment categorizes the group unit consisted of multiple containers with the same destination port, the same weight class and the same size, and the box unit. Recently, there are the studies dealt with containers at transshipment hubs as Lee *et al.* (2006) and Han *et al.* (2008). They do not consider the container flow at the terminal between the mega-containership and feeder ships. Then Nishimura *et al.* (2009) consider the problem for transshipment containers at the terminal where the mega-containership calls. They proposed the heuristic algorithm by using the subgradient method with lagrangian relaxation which is known by lower bound. Then in this paper, we propose meta-heuristics based algorithms and investigate the solution quality obtained by them, in order to develop the model by considering more real situation near future.

3. PROBLEM DEFINITION

The terminal where the mega-containership calls is not only used by the mega-containership and feeder ships but also by ships with other sizes. The yard area where the containers stored is divided roughly according to the direction of containers passing through that terminal such as inbound, outbound and transshipment. This paper deals with the problem for only transshipment containers related to mega-containerships. From our knowledge, the storage areas of all containers are divided into the directions of container movement into the terminal, there is less interference among them. Therefore we do not have to deal with the ships with other sizes in this step of our study.

First the target area of this problem is described. There is the container flows into a terminal in two ways: from the mega-containership to feeder ships, and from feeder ships to the mega-containership, as shown in Figure 1. Containers need to be stored in the yard for several reasons: time differences between ship arrivals, repositioning of containers in ship stowage and modifying the work sequence of container handling. In this study, to simplify the treatment of the problem, we consider the container flow in only one direction from the mega-containership to feeders, which the containers are stored in the yard area after the mega-containership arrives, and then those containers are stowed into the relevant feeder ship. Note that, in general, the loading operation to a ship cannot begin before the unloading operation to the relevant ship completes. This constraint has to be satisfied with in this paper. Therefore, if the unloading operation to a feeder ship has not completed yet when containers related to the relevant ship coming from the mega-containership have already been stored in the yard, beginning time of loading operation to this ship will be postponed until the unloading operation completes. On the contrary, if the containers coming from the mega-containership have not been stored in the yard when the unloading operation to a feeder ship completes, beginning time of loading operation to this

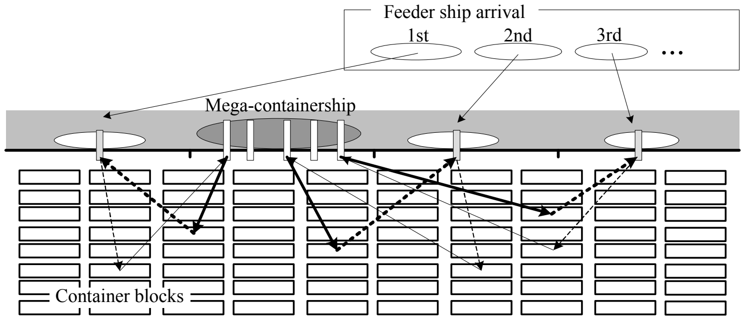


Figure 1. Container Flow in the Terminal

ship has to delay until containers are stored.

There are two movement types of transshipment containers as follows: (a) containers unloaded from one ship are stored at yard areas, and then those are loaded to the other ships, or (b) containers unloaded from one ship path through the yard areas, and those are loaded to the other ships directly. A very simple but not very correct way of considering this problem would be to assume that the containers move from the mega-containership to the feeder ships directly by simply transporting the containers through the yard area. In that case a constant time for handling operation between the mega-containership and feeder ships could be included in the objective function. Because the handling time spent by containers' movement between their ships would not depend on the way of container storage but also on the number of containers handled. However, from our knowledge, in order to rearrange the working sequence to ship stowage, most of containers are stored at the yard area and not just passing through. Therefore, in this study, we do not consider direct loading of containers to feeder ships. In general, the container stack arrangement and the container storage volume of the yard area are changing slowly during the relevant planning horizon. Our aim of this study is to formulate this problem in a simple way. Then, in future research, the implementation will be tested and further developed for more realistic applications. Thus, we consider the static model and compare the computational results obtained by multiple solution methods in this paper.

Secondly, we describe the treatment of containers handled. A yard area has multiple container blocks, and one container block consists of multiple containers. In this study, containers for a particular ship are treated as one container group. For example, one container group for "ship A", which consists of 100 containers, is stored in one container block in the yard. We do not assign a single container to a single location in the container block. Generally, if we should treat the problem in unit of a single container, we should have to deal with issues like:

– precedence constraints in the working sequence and stowage location for ship stowage, and also
– re-handle works caused by the working sequence and storage allocation in the yard,

by considering weight class and destination port for each container. However, as mentioned above, there is no need to deal with these issues in this study.

Thirdly, we state how to use a single block. Generally, multiple ships can share a specific block if the time of one ship serviced is different from that of other ships serviced. However Nishimura et al. (2009) assume that a part of the yard area in the target block is used for the containers handled of other ships in another planning period. If we assume that we consider this problem only with containers handled from the mega-containership during its stay, it is reasonable to assume that, containers which are loaded to one feeder ship are stored in single yard block, the yard block can be not shared by containers of multiple ships.

The determinant, the objective function and constraints of our problem are as follows:

the determinant is storage allocation for transshipment containers in the yard,

the objective function is total service times from the mega-containership to feeder ships including the waiting time for feeders until start of the loading operation,

constraints are

– each container group handled by a quay crane assigned to a feeder ship is stored exactly once in the yard,
– each container group for only one feeder ship is assigned to a container block in the yard, and
– the loading operation for feeders cannot begin until the containers coming from the mega-containership are stored in the yard.

4. PROBLEM FORMULATIONS

In this section, a mixed integer programming formulation of the container storage allocation problem (CSAP) and the lagrangian relaxation of this formulation are presented.

4.1 Formulation of the Container Storage Allocation Problem (CSAP)

In formulating the CSAP, a binary variable x_{jl} is defined to specify whether the container group for ship j is stored at the yard location l. We consider the CSAP formulation, which is modified from the formulation of the ship-to-berth allocation problem (BAP) in Imai *et al.* (2001). This formulation is further called [P-CSA]. Afterwards, we deal with berths, ships and order of service sequence in the BAP as quay cranes assigned to the mega-containership, feeder ships, an order of working sequence for quay cranes assigned to the mega-containership.

[P-CSA] Minimize

$$\sum_{i\in QC}\sum_{j\in V}\sum_{l\in YP}\{(T_i - K_j + 2)C_{B(i)jl}^{MY} + C_{B(j)jl}^{YF} + A_M - F_j\}Q_{ij}x_{jl} + \sum_{i\in QC}\sum_{j\in V}\sum_{l\in YP} Q_{ij}y_{jl} \quad (1)$$

subject to

$$\sum_{j\in V} x_{jl} = 1 \qquad \forall\, l \in YP, \quad (2)$$

$$\sum_{l\in YP} x_{jl} = 1 \qquad \forall\, j \in V, \quad (3)$$

$$\sum_{j'\in V}\sum_{l'\in YP} \frac{1}{K_j - K_{j'}} \max\{K_j - K_{j'}, 0\}(C_{B(i)j'l'}^{MY} x_{j'l'} + y_{j'l'})Q_{ij'} + Q_{ij}\{y_{jl} - (F_j - A_M - C_{B(i)jl}^{MY})x_{jl}\} \ge 0$$

$$\forall\, j \in V, l \in YP, i \in QC, \quad (4)$$

$$x_{jl} \in \{0, 1\} \qquad \forall\, j \in V, l \in YP, \quad (5)$$

$$y_{jl} \ge 0 \qquad \forall\, j \in V, l \in YP, \quad (6)$$

where

i (=1, …, NQ) $\in QC$ set of quay cranes
j (=1, …, NV) $\in V$ set of feeder ships
l (=1, …, NP) $\in YP$ set of container blocks in the yard
T_i number of ships handled by quay crane i assigned to the mega-containership
K_j order of unloading sequence at quay cranes assigned to the mega-containership for handling of feeder ship j's container group
Q_{ij} 1 if quay crane i handles the container group of feeder ship j, 0 otherwise
A_M arrival time of the mega-containership
F_j completion time of feeder ship j's containers for unloading
$C_{B(i)jl}^{MY}$ handling time spent by feeder ship j 's container movement from location of quay crane i assigned to the mega-containership $B(i)$ to yard block l
$C_{B(j)jl}^{YF}$ handling time spent by feeder ship j 's container movement from yard block l to ship j 's berthing position $B(j)$
x_{jl} decision variable; 1 if the container group of feeder ship j is stored in yard block l, 0 otherwise
y_{jl} stay time of feeder ship j 's container group in yard block l until beginning of loading operation to a relevant ship.

The objective function (1) minimizes the sum of handling time for the container movements from the mega-containership to yard blocks, and of handling time for the container movements from yard blocks to the feeder ships' berthing location, plus the waiting times for the feeder ships. Constraint set (2) ensures that the container group for each feeder ship must be stored at a yard block. Constraint set (3) enforces that each yard block services up to the container group for a feeder ship. Constraints (4) assume that the loading operation for a feeder ship must start after its arrival and its unloading operation. In the objective function, a handling time $C_{B(i)jl}^{MY}$ is weighted by $(T_i - K_j +2)$. $(T_i - K_j +1)C_{B(i)jl}^{MY}$, this results from the observation that the handling time $C_{B(i)jl}^{MY}$ of a specific ship j serviced by quay crane i assigned to the mega-containership contributes to the waiting time of the feeder ships, which are to be serviced by the same quay crane after it. In other words, the waiting time of a particular ship is represented by the cumulative handling time of its predecessors.

4.2 Lagrangian Relaxation of CSAP

The formulation of the CSAP is a mixed integer programming problem formulation which is not known to be solvable in polynomially-bounded time. In this section, we develop a heuristic procedure by the subgradient optimization procedure based on the following lagrangian relaxation of the original problem [P-CSA]:

[P-LR] Minimize

$$\sum_{i\in QC}\sum_{j\in V}\sum_{l\in YP}\{(T_i - K_j + 2)C_{B(i)jl}{}^{MY} + C_{B(j)jl}{}^{YF} + A_M - F_j\}Q_{ij}x_{jl} + \sum_{i\in QC}\sum_{j\in V}\sum_{l\in YP}Q_{ij}y_{jl}$$

$$-\sum_{i\in QC}\sum_{j\in V}\sum_{l\in YP}\lambda_{ijl}\sum_{j'\in V}\sum_{l'\in YP}\frac{1}{K_j - K_{j'}}\max\{K_j - K_{j'},0\}(C_{B(i)j'l'}{}^{MY}x_{j'l'} + y_{j'l'})Q_{ij'} - \sum_{i\in QC}\sum_{j\in V}\sum_{l\in YP}\lambda_{ijl}Q_{ij}\{y_{jl} - (F_j - A_M - C_{B(i)jl}{}^{MY})x_{jl}\} \quad (7)$$

subject to (2), (3), (5) and (6).

Where λ_{ijl} is the lagrangian multiplier for quay crane *i* assigned to the mega-containership, feeder ship *j*, and yard block *l* and has a non-negative value.

Because the y_{jl} 's do not appear in any constraint, they are redundant. By this, this formulation can be rewritten as follows:

[P1] Minimize

$$\sum_{i\in QC}\sum_{j\in V}\sum_{l\in YP}\{(T_i - K_j + 2)C_{B(i)jl}{}^{MY} + C_{B(j)jl}{}^{YF} + A_M - F_j\}Q_{ij}x_{jl}$$

$$-\sum_{i\in QC}\sum_{j\in V}\sum_{l\in YP}\lambda_{ijl}\sum_{j'\in V}\sum_{l'\in YP}\frac{1}{K_j - K_{j'}}\max\{K_j - K_{j'},0\}C_{B(i)j'l'}{}^{MY}Q_{ij'}x_{j'l'} + \sum_{i\in QC}\sum_{j\in V}\sum_{l\in YP}\lambda_{ijl}(F_j - A_M - C_{B(i)jl}{}^{MY})Q_{ij}x_{jl} \quad (8)$$

subject to (2), (3) and (5).

Problem [P1] is further reformulated by introducing the representative cost E_{jl} in the objective function.

[P2] Minimize

$$\sum_{j\in V}\sum_{l\in YP}E_{jl}x_{jl} \quad (9)$$

subject to (2), (3) and (5).

By relaxing constraint set (4), formulation [P-CSA] becomes a two-dimensional assignment problem and is therefore easy to solve.

5. SOLUTION PROCEDURE

5.1 Subgradient method

The quality of the feasible solution obtained through the above procedure strongly depends on the ability to determine good lagrangian multipliers λ_{ijl}. In [P-LR], for each combination of (*i*, *j*, *l*), it is assumed that the container group for each ship can be allocated to yard blocks while minimizing total handling time from the mega-containership to feeder ships and total waiting time for feeder ships. This may lead to an infeasible solution of the original problem [P-CSA] as the container groups for some feeder ships may be loaded before these ships have completed their unloading operation. Good multipliers are also important as the quality of the lower bound, (i.e., the objective function value of [P-LR]), is a function of those multipliers. The best lower bound corresponding to the optimal multiplier vector λ^* is determined as

$$Z_{LR}(\lambda^*) = \text{Max}_{\lambda}(Z_{LR}(\lambda)), \quad (10)$$

where $Z_{LR}(\lambda)$ is the value of the lagrangian function with a multiplier set (vector) λ. To find the set of the lagrangian multipliers, we have used the subgradient optimization procedure (Fisher, 1981). This approach has been widely utilized. It is sufficient to note that as a termination criterion the maximum number of iterations has been fixed at 200. The procedure also terminates if the gap between the feasible solution value and the lagrangian bound becomes less than 1. Given integer objective function coefficients, this condition is sufficient to detect an optimal solution. The process, discussed in the next section, is performed at each iteration of the heuristic to determine a feasible solution to [P-CSA]. At the time of termination, the subgradient optimization procedure reports the best feasible solution and the best lower bound generated in all the iterations.

During each iteration of the procedure, [P-LR] is solved to obtain a lower bound for [P-CSA]. Note that the objective function value of [P-LR] is equal to the objective function value of [P2] plus the value related to y_{jl}. As the objective function of [P-LR] is a lower bound, y_{jl} values are fixed to zero. The subgradient procedure is detailed in Nishimura *et al.* (2009). We propose the following process to find the feasible solution. The indices *i* and *j* indicate the quay crane number and the ship number, respectively.

Step 1 : Let $i = 1$, $j = 1$, $Z_{Feas} = 0$.

Step 2 : If $j >$ the number of ships, Let $i = i + 1$, $j = 1$ and $FIRST = 0$. If $i >$ the number of quay cranes assigned to the mega-containership, STOP.

Step 3 : If $Q_{ij} = 1$ (i.e., ship *j* is serviced by quay crane *i*), go to step 4. Otherwise go to step 8.

Step 4 : Get CT^{MY} ; the completion time of ship *j* 's containers unloading from quay crane *i* to yard location *l* as following: If $FIRST = 0$, $CT^{MY} = A_j + C_{B(i)jl}{}^{MY}$. Otherwise, $CT^{MY} = CT^{MY} + C_{B(i)jl}{}^{MY}$. And let $FIRST = 1$.

Step 5 : Get BT_j^{YF} ; the beginning time of ship j 's containers loading from yard block l to feeder ship j 's berthing position as following: If $A_j < CT^{MY}$, $BT_j^{MY} = CT^{MY}$ and $y_{jl} = 0$. Otherwise, $BT_j^{MY} = A_j$ and $y_{jl} = A_j - CT^{MY}$.
Step 6 : Get CT_j^{YF} ; the completion time of ship j 's containers loading from yard block l to feeder ship j 's berthing position as following: $CT_j^{YF} = BT_j^{FY} + C_{B(j)jl}^{MY}$.
Step 7 : Calculate the objective function of a feasible solution, $Z_{\mathrm{Feas}} = Z_{\mathrm{Feas}} + C_{B(i)jl}^{MY} + CT_j^{YF} - A_j$.
Step 8 : Let $j = j + 1$ and go to step 2.

5.2 GA based heuristic

As stated in Zhang *et al.* (2001a), a group of chromosomes constitutes a population. A merit is assigned to each chromosome, depends on fitness value. A new generation is evolved by a selection technique, in which there is larger probability of the fittest chromosomes being chosen. Pairs of chosen chromosomes are used as the parents in the construction of the next generation. A new generation is produced as a result of reproduction operators applied on crossover. New generations are repeatedly produced until a predefined convergence level is reached.

We explain the chromosome representation which can be expressed as the candidate of solution. Then, each container group is stored at only one yard block. In Figure 2, we assume that ten yard blocks are prepared as stack locations for ten container groups. As explained in Sections 4 and 5, the container group for each feeder ship must be stored at a yard block, and each yard block services up to the container group for a feeder ship. Thus in this example, container groups 7, 2, 10, 3, 4, 1, 9, 5, 8 and 6 are stored at yard blocks 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10, respectively. We consider the representation to allow each weighting to range between 1 and *n* (# of container groups). In this paper, chromosomes may be variable in length, and the corresponding candidate solution may be underspecified (a given gene is not represented) or over-specified (a single gene is represented more than once). In GA heuristics of this paper, locus and gene are expressed by container group number and yard block number, respectively.

We explain the selection how to determine the pair of chromosome, the crossover how to generate new chromosomes, the mutation how to arrange the chromosomes as follows.

Selection Mechanisms: tournament selection.

Crossover Operation: partially matched crossover

Mutation: In order to determine whether a chromosome will undergo a mutation operation, a random test is performed. The random selection test is based on generating a random number *p* (=[0, 1]) for the considered chromosome. If a number *p* for a chromosome is larger than the mutation rate predetermined, the chromosome will undergo mutation. In mutation operation, only one locus (i.e., container group number) is selected randomly, the gene (i.e., yard block number) of a relevant container substitutes another gene.

GA Parameters: after the preliminary experiments are conducted, generations which terminations means in GA are determined to 5000, the number of individuals in a generation is set to 50 for our heuristic.

Yard block #	1	2	3	4	5	6	7	8	9	10
Container group # related to feeder ship *j*	7	2	10	3	4	1	9	5	8	6

Figure 2. Chromosome representation

Yard block #	1	2	3	4	5	6	7	8	9	10
Candidate solution	7	2	**10**	3	4	1	9	**5**	8	6
New solution	7	2	**5**	3	4	1	9	**10**	8	6

Figure 9. Swap move for tabu search

5.3 Algorithm by using TS

The fundamental mechanism of TS is that potentially reversing the effects of previous moves by interchanges that might return to previous positions, in order to prevent the search from repeating swap combinations tried in the recent past. The memory structures in TS operate by reference to following dimension consisting of recency, frequency, quality and influence. The most commonly used short term memory keeps track of solutions attributes that have changed during the recent past, and is called recency-based memory or tabu list. To exploit this memory, selected attributes that occur in solutions recently visited are labeled tabu-active, and solutions that contain tabu-active elements, or particular combinations of these attributes, are those that become tabu. Frequency-based memory provides a type of information that complements the information provided by recency-based memory, broadening the foundation for selecting preferred moves. Quality memory refers to the ability to differentiate the merit of solutions visited during the search. TS concept of quality is broader than the one implicitly used by standard optimization methods. Influence memory considers the impact of the choices made during the search, not only on quality but also on structure. Recording information about the influence of choices on particular solution elements incorporates an additional level of learning.

The flow of TS is as follows. First of all, we generate the initial solutions *x_now*, calculate the objective function value f(*x_now*) of them, finds better solution with shortest service time from them, it is set as best solution *x_best* at the current iteration. Next, we generate multiple solutions *x_next*(s) from *x_now*, and then add move attribute of *x_next*(s) to diversification-memory. If given penalty parameter related to diversification is less than the number of move occurring *x_next*(s), introduce penalty to *x_next*(s). Find better solution *x_better* from *x_next*(s). And add move attribute of *x_better* to intensification-memory. If given penalty parameter related to intensification is less than the number of move occurring *x_better*, introduce penalty to *x_better*. If move attribute of *x_better* is in tabu list and f(*x_best*) is greater than f(*x_better*), *x_best* is set to *x_better* and update tabu list. If f(*x_now*) is greater than f(*x_better*), *x_now* is set to *x_better*. If current iteration is less than the termination, continue the process. Otherwise, the process is ended.

As the solution representation in TS, We use the chromosome representation proposed in GA as shown in Figure 6 this concept is as well as that of GA. Move attribute in this problem is as follows. Firstly, we determine two locations randomly from the current solution, and exchange both locations as shown in Figure 9, generate new solution. Aspiration criteria are introduced in TS to determine when tabu restrictions can be overridden, thus removing a tabu classification otherwise applied to a move. The appropriate use of such criteria can be very important for enabling a TS to achieve its best performance levels. Aspiration criteria employed in this paper is a simple type consisting of removing a tabu classification from a trial move when the move yields a solution better that the best obtained until current iteration. The recency based memory respectively identify the starting and ending iterations of the tabu tenure for specified attribute " *e* ", thus bracketing the period during which " *e* " is tabu-active. This tabu tenure is called as tabu list size in most of related papers. In this paper, tabu list size is set to 7. The number of candidate solutions generated from the current solution in a iteration is set to 60, and the termination is set to 2000.

6. NUMERICAL EXPERIMENTS

6.1 Experimental design

The solution procedure is coded in the "C" language. Problems used in the experiments are generated randomly, but systematically. We consider the terminal layout shown in Figure 10. The mega-containership is serviced at berth 2 in linear type and done at berths 2 and 3 in indented type. We assume that the average intervals of feeders' completion time for unloading are 1 hour and 2 hours. If we consider the problem for a long time period, the arrival time of the mega-containership may not be so important.

The arrival time of the mega-containership effects on the number of feeder ship waiting and the waiting time length of feeder ship. For instance, comparing early arrival of the mega-containership with its late arrival against feeder ships' arrival, ships waiting in the latter one is more than those in the former one. From this reason, the arrival time of the mega-containership effects on the objective function value. In order to investigate how large the difference among the arrival times of the mega-containership, we assume that A_M has a unique value in our experiments, we performed four computations with different values for A_M. In our experiments A_M was set to 1/2 and 3/4 of time duration between the first and final feeder ships' completion time of unloading operation, and final feeder ships' completion time of unloading operation, indexed by time 1, 2 and 3.

In this problem, it is assumed that only one quay crane is assigned to each feeder ship. For this reason, between 50 and 250 containers are stowed to each feeder ship. Although the maximum number of containers flow from the mega-containership to feeder ships might be up to 10,000, we assume that about half of 10,000 containers are dealt with. Thus the mega-containership calling once treats the containers for about thirty feeder ships. With respect to the handling time for each feeder ship's container from the mega-containership to the yard block, we assume that it takes 1.2 minutes to handle a container at the yard blocks closest to either the locations of the quay crane assigned to the mega-containership or the feeder ships' berthing location. This time is obtained by the following formulation with the number of containers handled and the distances from yard blocks to the location where a relevant feeder ship is berthing:

6.2 Computational results

In this section, we show the computational results under parameters we stated above.
The total service times as objective function value are shown in Table 1, those are the average values among twenty cases combined five types of completion times of feeder ships' loading and five types of handling times between yard blocks and mega-containership or feeder ships. Note that AM means the mega-containership arrives, we call time-1 of the mega-containership arrival as AM-1.

In order to compare among the total service times by three algorithms, we analyze the results as following by each time of the mega-containership arrival:
- At AM-1 (mega-containership early arrival); at 1 hour average interval of completion time for feeders' unloading (hereafter, 1 hour interval), best algorithm depends on the terminal layout. In the linear type, the total service time obtained by LR is shorter than that time done by others. In the indented type, that time obtained by LR is shorter that time done by others. At 2 hours interval, there is not difference between the total service time obtained by GA and that time done by TS. Those times obtained by them are shorter than that done by LR. At 3 hours interval, the total service time obtained by TS is shorter than that done by others. Thus the longer interval of containers' arriving at the yard, the shorter the total service times obtained by TS are.
- At AM-2; best algorithm with shortest total service time depends on the terminal layout. In the linear type, in each average interval of completion times for feeders' unloading, the total service time obtained by LR is better than that done by others. In the indented type, best algorithm depends on average interval of completion time for feeders' unloading. At 1 hour interval, the total service time obtained by LR is better than that done by others. At 2 hours interval, there is not difference among the total service time obtained by all algorithms. At 3 hours interval, the total service time obtained by GA is shorter than that done by others. Thus in this situation, best algorithm with shorter total service time depends on the terminal layout and the congestion rate of the terminal.
- AM-3 (mega-containership lately arrival); at both terminal layouts and all average interval of completion time for feeders' unloading, those obtained by LR are smaller than those done by others. This reason is that the solutions obtained by LR are the optimal solutions as shown in Table 2.

As shown in Table 1, from the average objective function value only, we cannot know the better algorithm at each case. Thus secondly, in order to investigate the results by individual problem in more detail, we count the best algorithm with shortest total service time shown in Table 2. Note that, if there are best solutions in two or more algorithms, we include them as best one in a relevant algorithm. Thus total number of best solutions does not equal to twenty five. ( ) means the number of cases obtained the optimal solution. We investigate largest number of best solutions obtained, in most cases of the best algorithms shown in Table 1, the algorithms with shorter total service times become those with larger number of best solutions. Then in order to investigate comparison among algorithms more detail, Table 2 includes the ranking of great number of best solutions among algorithms.

In order to compare among the number of best solutions with shorter total service times by three algorithms, we analyze the results as following by each time of the mega-containership arrival:
- At AM-1 (mega-containership early arrival); at 1 hour interval, best algorithm depends on the terminal layout. We put the best solutions obtained by each algorithm in decreasing order. LR, TS and GA are in order of better algorithm in the linear type. LR, GA and TS are in order of better algorithm in the indented type. At 2 and 3 hours interval, best algorithm also depends on the terminal layout. TS, GA and LR are in order of better algorithm in the linear type. GA, TS and LR are in order of better algorithm in the indented type. Therefore, comparing with GA and TS, TS is better than GA in the linear type, GA is better than TS in the indented type.- At AM-2; at 1 hour interval, best algorithm depends on the terminal layout. We put the best solutions obtained by each algorithm in decreasing order. As well as AM-1, LR, TS and GA are in order of better algorithm in the linear type. LR, GA and TS are in order of better algorithm in the indented type. At 2 hour interval, best algorithm does not depend on the terminal layout. Putting the best solutions obtained by each algorithm in decreasing order, as well as the result of 1 hour interval and linear type, LR, TS and GA are in order of better algorithm at both layouts.
- At AM-3 (mega-containership lately arrival); best algorithm depends on the terminal layout, does not depend on average interval of completion time for feeders loading. LR, TS and GA are in order of better algorithm in the linear type. LR, GA and TS are in order of better algorithm in the indented type.
Therefore, best algorithm is LR. And comparing with GA and TS, as well at AM-1, TS is better than GA in the linear type, GA is better than TS in the indented type.

Table 1. Total service time (average value among all cases (hours))

	Average interval of CTFU	Linear type LR	GA	TS	Best	Indented type LR	GA	TS	Best
AM-1	1 hour	587.9	589.3	588.5	LR	426.2	425.8	425.9	GA
	2 hours	566.6	564.2	564.2	GA,TS	467.2	464.8	464.8	GA,TS
	3 hours	613.1	610.1	609.9	TS	525.1	521.7	521.6	TS
AM-2	1 hour	779.1	780.8	779.7	LR	606.9	607.2	607.5	LR
	2 hours	925.0	926.4	925.5	LR	760.8	760.8	760.8	Even
	3 hours	1071.4	1072.3	1072.1	LR	931.7	930.9	931.0	GA
AM-3	1 hour	991.4	993.1	992.0	LR	819.2	819.5	819.9	LR
	2 hours	1351.2	1352.8	1351.7	LR	1178.9	1179.3	1179.5	LR
	3 hours	1711.3	1712.9	1711.8	LR	1539.1	1539.4	1539.5	LR

* CTFL means completion time for feeders' unloading

Table 2 Number of best solutions with shortest total service time (hours)

	Average interval of CTFU	Linear type LR	GA	TS	Decreasing.	Indented type LR	GA	TS	Decreasing.
AM-1	1 hour	25(25)	2(2)	5(5)	LR>TS>GA	13(3)	13(1)	7	LR>GA>TS
	2 hours	6	9	11	TS>GA>LR	4	14	12	GA>TS>LR
	3 hours	5	8	15	TS>GA>LR	0	17	13	GA>TS>LR
AM-2	1 hour	25(25)	0	5(5)	LR>TS>GA	25(25)	11(11)	7(7)	LR>GA>TS
	2 hours	25(25)	1(1)	5(5)	LR>TS>GA	13(3)	9(2)	11	LR>TS>GA
	3 hours	23(14)	6(3)	5(2)	LR>GA>TS	10	16	9	GA>LR>TS
AM-3	1 hour	25(25)	0	5(5)	LR>TS>GA	25(25)	10(10)	6(6)	LR>GA>TS
	2 hours	25(25)	0	5(5)	LR>TS>GA	25(25)	10(10)	2(2)	LR>GA>TS
	3 hours	25(25)	0	5(5)	LR>TS>GA	25(25)	10(10)	5(5)	LR>GA>TS

* CTFL means completion time for feeders' unloading
( ) means the number of optimal solutions obtained

Thirdly, LR can obtain the lower bound and feasible solution simultaneously. Thus we will compare the lower bound with the value of feasible solutions, in order to investigate the quality of solutions obtained by each algorithm. Table 3 shows the difference between the lower bound and the value of feasible solution as gap (%) expressed in following equation (16).

$$\text{Gap}(\%) = \frac{\text{feasible solution value} - \text{lower bound}}{\text{lower bound}} \times 100 \quad (16)$$

Note that, the value 0.00 means that the optimal solution is obtained. At AM-3 when the mega-containership arrives after all feeder ships arrival, gap (%) is 0.00 in both layouts and all congestion of feeder arrival. Note that the algorithm which has smaller gap is better algorithm. In order to compare among the number of best solutions with gap by three algorithms, we analyze the results as following by each time of the mega-containership arrival:

In the linear type, the algorithms with smaller gap shown in Table 3 are same as the algorithms with large number of best solutions shown in Table 2. There is the same tendency between the algorithm with smaller gap and the algorithm with number of best solutions. In the indented type, there is not the same tendency between them.

- At AM-1 (mega-containership early arrival); at 1 hour interval, gap depends on the average interval of completion time for feeders' unloading. We put the best solutions obtained by each algorithm in increasing order. GA, TS and LR are in order of better algorithm. At 2 and 3 hours interval, TS, GA and LR are in order of better algorithm.
- At AM-2; at 1 hour interval, gap also depends on the average interval of completion time for feeders' unloading. We put the best solutions obtained by each algorithm in increasing order. LR, GA and TS are in order of better algorithm. At 2 hour interval, there is not difference between gap obtained by GA and TS. Gaps obtained by them are smaller than that done by LR. At 3 hour interval, GA, TS and LR are in order of better algorithm. According to the case, best algorithm is various.
- AM-3 (mega-containership lately arrival); best algorithm does not depend on average interval of completion time for feeders loading. LR, GA and TS are in order of better algorithm.

Additionally, the service times for the mega-containership are shown in Table 4. In all situations, those times obtained by TS are longer than those done by others. At AM-2 and AM-3 or 1 hour interval, the service times for the

Table 3. Gap (%) between lower bound and objective function value of feasible solution

	Average interval of CTFU	Linear type				Indented type			
		LR	GA	TS	Increasing.	LR	GA	TS	Increasing.
AM-1	1 hour	0.00	0.23	0.10	LR<TS<GA	3 03	2.90	2.94	GA<TS<LR
	2 hours	7.10	6.43	6.40	TS<GA<LR	42. 29	41.12	41.06	TS<GA<LR
	3 hours	44.31	43.03	42.92	TS<GA<LR	236. 67	232.95	232.67	TS<GA<LR
	Average interval of CTFU	Linear type				Indented type			
		LR	GA	TS	Increasing.	LR	GA	TS	Increasing.
AM-2	1 hour	0.00	0.21	0.07	LR<TS<GA	0 00	0.04	0.10	LR<GA<TS
	2 hours	0.00	0.16	0.06	LR<TS<GA	1 27	1.26	1.26	GA,TS<LR
	3 hours	0.12	0.21	0.18	LR<TS<GA	4 44	4.33	4.34	GA<TS<LR
	Average interval of CTFU	Linear type				Indented type			
		LR	GA	TS	Increasing.	LR	GA	TS	Increasing.
AM-3	1 hour	0.00	0.17	0.06	LR<TS<GA	0 00	0.04	0.08	LR<GA<TS
	2 hours	0.00	0.12	0.04	LR<TS<GA	0 00	0.03	0.05	LR<GA<TS
	3 hours	0.00	0.10	0.03	LR<TS<GA	0 00	0.02	0.03	LR<GA<TS

* CTFL means completion time for feeders' unloading

Table 4 Service time for mega-containership (average value among all cases (hours))

	Average interval of CTFU	Linear type				Indent type			
		LR	GA	TS	Best	LR	GA	TS	Best
AM-1	1 hour	25.5	25.3	28.6	GA	16.7	16.6	17.8	GA
	2 hours	25.8	26.4	28.4	LR	16.7	17.6	19.0	LR
	3 hours	26.8	27.5	28.9	LR	16.7	17.3	18.9	LR
	Average interval of CTFU	Linear type				Indent type			
		LR	GA	TS	Best	LR	GA	TS	Best
AM-2	1 hour	25.5	25.2	28.6	GA	16.6	16.5	18.3	GA
	2 hours	25.5	25.3	28.7	GA	16.7	16.6	18.2	GA
	3 hours	25.5	25.4	28.8	GA	16.7	16.6	19.1	GA
	Average interval of CTFU	Linear type				Indent type			
		LR	GA	TS	Best	LR	GA	TS	Best
AM-3	1 hour	25.5	25.2	29.0	GA	16.6	16.5	18.5	GA
	2 hours	25.5	25.2	29.5	GA	16.6	16.5	18.3	GA
	3 hours	25.5	25.2	28.0	GA	16.6	16.5	18.7	GA

* CTFL means completion time for feeders' unloading

mega-containership done by GA are shorter than those done by others. On the contrary, those done by TS is longer than those done by others in all cases. For the total service time shown in Table 1, when the best algorithm is TS, LR becomes best algorithm in Table 4. When the best algorithm is LR or GA, GA becomes best algorithms in Table 4. The service times for the mega-containership depend on the terminal layout. It takes around 25 to 30 hours in the linear type, around 16 to 19 hours in the indented type. This reason is that there is the difference between the number of quay cranes assigned to the mega-containership in the linear type and indented type.

7. CONCLUSIONS

The port where the mega-containerships call, plays an important role in the connection between a relevant ship and others. This paper addresses the storage planning problem of transshipment containers which a containers which are dealt with at the mega-port, and suggests a solution procedure. The contribution of this paper to the literature is that we consider the storage problem at the terminal where the mega-containership calls under a container flow which is different from the previous research, and compare the algorithm by previous research with the algorithm by meta-heuristics: genetic algorithm (GA) and tabu search (TS). We obtain the total service time for the mega-containership and feeder ships as the objective function from the sub-gradient method by using lagrangian relaxation (LR), GA and TS, and the service time for only mega-containership, simultaneously. As the results from the numerical experiments, the earlier the mega-containership arrives after most of feeder ships arrive, the better the solutions obtained by LR. If the mega-containership arrives after a half of feeder ships arrive, TS is better than others in the linear type, GA is better than others in the indented type. Best algorithm depends on the arrival time of the mega- containership or the terminal layout. From the service time for the mega-containership, GA is better than others in most of cases. Our proposed algorithm in this paper will be able to modify the model and use that, in order to consider more realistic situation near future.

ACKNOWLEDGEMENTS

This research is partially supported by the MEXT under the Grant-in-Aid for Young Scientists-B Grant 21710453.

8. REFERENCES

Bazzazi, M., Safaei, N., Javadian, N. (2009). A genetic algorithm to solve the storage space allocation problem in a container terminal. *Computers & Industrial Engineering* 36, 1711-1725.

M.L. Fisher . (1981). The Lagrangian relaxation method for solving integer programming problems, *Management Science* 27, 1–18.

Han, Y., Lee, L.H., Chew, E.P., Tan, K.C. (2008) A yard storage for minimizing traffic congestion in a marine container transshipment hub, *OR Spectrum* 30(4), 679-720.

Kim, H.K., Hong, G.P. (2006). A heuristic rule for relocating blocks, *Computers & Operations Research* 33(4), 940-954.

Kim, K.H., Park, Y.M., Ryu, K.R. (2000). Deriving decision rules to locate export containers in container yards. *European Journal of Operational Research* 124, 89-101.

Kim, K.H., Park, K.T. (2003). A note on a dynamic space-allocation method for outbound containers. *European Journal of Operational Research* 148, 92-101.

Kozan, E., Preston, P. (1999). Genetic algorithms to scheduling container transfers at multimodal terminals. *International Transactions in Operational Research* 6, 311-329.

Lee, L.H., Chew, E.P., Tan, K.C., Han, Y. (2006). An optimization model for storage yard management in transshipment hubs, *OR Spectrum* 28(4), 539-561.

Lim, A., Xu, Z. (2006). A critical-shaking neighborhood search for the yard allocation problem, *European Journal of Operational Research* 174, 1247-1259.

Nishimura, E. Imai, A., Janssens, G.K., Papadimitriou, S. (2009). Container storage and transshipment marine terminals, *Transportation Research Part E* 45(5), 771-786.

Park, C., Seo, J. (2009). Mathematical modeling and solving procedure of the planar storage location assignment problem, *Computers & Industrial Engineering* 57(3), 1062-1071.

Park, C., Seo, J. (2010). Comparing heuristic algorithms of the planar storage location assignment problem, *Transportation Research Part E* 46(1), 171-185.

Yang, J.H., Kim, K.H. (2006). A grouped storage method for minimizing relocations in block stacking systems, *Journal of Intelligent Manufacturing* 17(4), 453-463.

Zhang, C., Liu, J., Wan, Y.W., Murty, K.G., Linn, R.J. (2003). Storage space allocation in container terminals. *Transportation Research Part B* 37, 883-903.

Session I3: Seaport and Transportation 9

·**Day2: Sep. 16 (Thu.)**

·**Time: 16:20 - 17:40**

·**Chair: Axel Hahn**

·**Room: Iris, 4F**

CONTAINER LOGISTICS AND DATA MINING

Robert Stahlbock[1*] and Stefan Voß[1]

[1]University of Hamburg
Germany
e-mail: stahlbock@econ.uni-hamburg.de
stefan.voss@uni-hamburg.de

Abstract: We consider the application of data mining techniques for gaining insight into processes and problems of container handling. Container terminals perform similar functions, but the processes, technology, and labor requirements differ at each terminal. Therefore, terminals are faced with different bottlenecks. Due to the complexity of terminal operations it is difficult to identify bottlenecks within a process. The large amount of transaction data and the large number of potential, interdependent factors make an analysis challenging. However, having a large database is a prerequisite for using promising and well established data mining methods, such as decision trees, neural networks, support vector machines, or association rule analysis in order to derive knowledge from data. Data mining methods and algorithms are useful for analysing current data reflecting current processes, identifying problems, interpreting solutions as well as for forecasting. For example, forecasting the movements of full and empty containers on a terminal-wise, regional or even more aggregated level can be helpful for equipment planning purposes, since for liner operators the empty container repositioning problem is a very important issue. Simple decision trees can help to interpret data and identify causes for inefficiently performed processes (e.g., abnormal high truck dwell times within a terminal). Although there is a lot of scientific literature on either container logistics or data mining, there are only a very few publications which combine both research fields. This paper aims at initiating further research in this fruitful field by discussing possibilities for applying data mining techniques for solving container handling problems and supporting decision makers.

VIRTUAL TERMINAL OPERATION SYSTEM FOR JOINT EVALUATION OF OPERATION STRATEGIES

Axel Hahn

Business Engineering
University of Oldenburg
Ammerländer Heerstr. 118
26111 Oldenburg, Germany
email: hahn@wi-ol.de

1. INTRODUCTION

The last decade saw a considerable growth in worldwide container transportation and with it an indispensable need for optimization. Even with shrinking capacity utilization the terminal operations have to be cost optimal to face severe competition on international markets. In addition complexity of harbour operations is rising due to new requirements (security) and value added logistic services. Physical terminal layout, container handling technologies and operation software have to be jointly optimized. Due to their complexity the decomposition of planning and scheduling problems is a well-known technique to make these problems tractable. The resulting dependable sub problems were often optimized and managed using different planning systems (Meier and Schumann 2007). (Stahlbock and Voß 2007) identifies the need of integration of different planning approaches and tools. Latest at run time planning systems are jointly embedded in a Terminal Operation System (TOS) and cooperatively run the terminal. This paper introduces a generic framework for pre ante test of terminal operation software components in a framework which includes a physical simulation engine and virtual reality visualization.

2. TERMINAL OPERATIONS AND OPTIMIZATION

Optimization methods for container terminals are well analysed and described today (Stahlbock and Voß 2007). Simulation is an usual approach to evaluate and to select algorithms or heuristics for operation planning. Port simulation tools are available on the market and are established methods to optimize terminal operation (Cantrella, Cartenì, and de Luca 2005). By using agent platforms several researchers have also analysed the required joint application of different operation policies for subsystems (Henesey, Davidsson, and Persson 2008) and the coordination of interdependent planning systems. Terminal Operation Systems (TOS) integrate operative planning mechanisms in a joint operative software system. The efficiency of the terminal and the power of a TOS are depending on the cooperation of the combined planning systems (yard, crane, berth, ATV movements etc.). TOS system provider and user need tools to analyse the combined usage of different planning systems with respect to the actual port situation as identified by (Stahlbock and Voß 2007). Therefore performance analysis of a TOS has to be done in the context of the terminal it is intended for.

The company TBA has developed a methodology using advanced simulation models to test a TOS before it is actually deployed . This helps users to select and to verify its functionality. Approaches like from TBA help to evaluate the TOS as a complex software system. A flexible way to experiment with different planning heuristics and their cooperation is missing. Agent based simulation approaches have the power to evaluate the different optimization and planning mechanisms (Henesey, Davidsson, and Persson 2008)(Henesey 2006). Approaches like (Franz, Stefan Voss, and Rölke 2007) additionally integrate market aspects. The deployment, integration and joined analysis of different planning mechanisms for productive usage inside a TOS-System cannot be analysed by existing agent systems because of the different system architectures and missing integration in the usage environment. Today industrial simulation tools like Flexsim (http://www.flexsim.com/) provide 3D rendering as a standard functionality. Compared to 3D visualisation Virtual Reality allows a better immersion of the user. Interactive and immersive presentation of the port model improves understanding and better space recognition (Ryan 2001). VR technology needs an efficient authoring and modelling process chain to keep costs low.

3. VIRTUAL TESTBED FOR TERMINAL OPERATION SYSTEMS TERMINAL OPERATIONS

vTOS will be a TOS test bed within the terminal simulation environment Virtual Port which will be introduced in this contribution. Virtual Port is a generic environment to integrate terminal operation software, physical simulation engines and immersive real time virtual reality visualization. vTOS is an open and generic framework of a TOS that allows the integration of different planning mechanisms to evaluate their cooperation and interoperability in a production near software system.

The architecture of the generic framework is sketched in figure 1. Interfaces link terminal operation software, physical simulation engines and virtual reality visualization. The framework provides loose system coupling for synchronous and asynchronous simulation and demonstration. vTOS is the platform basis for the virtual and experimental terminal operating system. As already mentioned a simulation software (named SIM) and a Virtual Reality system (named VIS) is added to this scenario. The vTOS-DB stores the data required for the port operation. The interface between the vTOS and SIM is a service based equivalent to the real interfaces of a TOS to its operative entities (ATV, Bridges, etc.). The SIM is implemented by Plant Simulation from Siemens and provides a physical simulation of all port operations. Behaviour of ATV or manually operated transportation systems are implemented within this system.

We added a service based interface (Jammes and Smit 2005) to Plant Simulation which is alike real services to address and to communicate with the terminal operation systems.

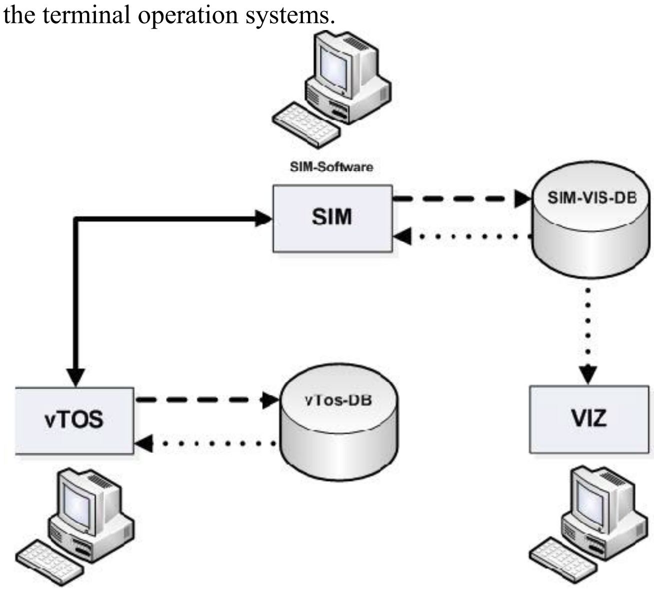


Figure 1. vTOS Architecture

The physical movements of all entities are logged in a database names SIM-VIS-DB. This database can be read for synchronous of asynchronous visualization by VIZ a 3D virtual reality model and rendering system. Synchronous visualization allows interaction with the user to manipulate the behaviour of the entities. This information has impact on the model simulated within SIM. Asynchronous visualization allows rewind, slow down and fast forward the time line for detailed inspection.

Each object within SIM has a 3D model which is in the Database to be loaded, visualized and moved by VIZ. For the Virtual Reality rendering the Xinth3D engine is used. For each simulation cycle an update of the positions of all entities are loaded from the database and the trajectories for all movements are calculated.The use of the Xinth3D engine provides significant higher performance, 3D stereo rendering and better quality the visualization module within Plant Simulation.

The framework can be used for evaluation, simulation, verification and – due to its 3D virtual reality interface – for training and education. This has already been done for planning systems for berth allocation, ATV and bridge scheduling as a proof of concept with the Jade Weser Port which is under construction at Wilhelmshaven, Germany (http://www.jadeweserport.com/). The concept is evaluated by implementing and testing methods for berth allocation, crane scheduling, vehicle dispatching, a storage location. For berth allocation the simulated annealing concepts of (Kim and Moon 2003) and of Guan and Cheung in (Günther and Kim 2005) by tree search and pair exchange are implemented. For Crane Scheduling the branch and bound method of (Peterkofsky and Daganzo 1990) and the spacial constraints method of (Zhu and Lim 2004) are implemented. For vehicle dispatching the evolutionary algorithms described in (Böse et. al. 2000) is used. For Yard Planning (Storage Location Problem) the approach of (Zhang et. al. 2003) is implemented. vTOS is able to apply and exchange this methods for evaluation for a specific terminal.

4. SERVICE INTEGRATION

As a generic platform vTOS offers an interface to embed planning tools and uses a service based interface (Bieberstein 2008) to the terminal operation entities simulated by objects in Plant Simulation.

Plant Sim ulation is an object orien ted even t driven (Banks 2005) si mulation t ool and serv ice requests are dispatched to the objects. Local in telligence (e.g. for routing of a Van Carrier done by the driver) is m odeled with the class (defining the Van Carrier behavior and physical parameters) in Plant Simulation.

The Communication between the vTOS and the Simulation tool is implemented by a set of messages transported by two communication channels (sockets). The messages from TOS to the terminal are

1. VC Container Movement: [TaskID, ID, VC ID, Origin, Destination, ContainerID]
   Message to (Van) Carrier etc. to move container. Location can be on train, truck, dock or yard.
2. Ship Movement: [TaskID, ID, ShipID, Length, Destination]
   Destination is the geoinformation of the bow of the vessel.
3. Train Movment: [TaskID, ID, TrainID, Length, Destination]
   Length is the number of wagons
4. PlantsSimulation Control: [TaskID, ID, Event]
   This is the only message which is no transport task. Main events are stop and reset.
5. Container bridge-Crane Movement: [TaskID, ID, CranID, Destination]
6. Container bridge-Container Movement: [TaskID, ID, CraneID, Type, Destination, ContainerID]
   Loading/unloading operation
7. Transtainer – Movement: [TaskID, ID, TranstainerID, Destination]
   Like container bridges the transtainer will be ordered at a specific position at the train rails.

Additional there are messages for the back channel and for new tasks to be planned.

At SIM (Plant Simulation) a method receives the messages and dispatches them to the target objects representing a dedicated entity. The obj ects/entities handle their job autonomous, tracing their movements in th e SIM-VSD-database and return status information to vTOS.

The representation is quite static and the user is not able to get an impression about space, vehicle speed etc. Additionally an interaction with entities would be more natural with its 3D representation. Therefore the traces of all entities movements are captured and can be read by the VIS component for VR representation.

During runtime the 3D engine accesses the traces stored in the SIM-VIS-DB synchronously (using system time) or asynchronously. The synchronous mode allows interaction with the system which will go back the way to the vTOS if necessary. The asynchronous mode allows a better analys is over tim e (rewind, fast forward, etc.). T he data schema model of t he SIM-VIS-DB is depi cted i n Fi gure 2. Seven m ain dat abase t ables store t he i nformation ab out the infrastructure, and all moving entities in addition to path information generated by the behaviour of the entities in SIM.

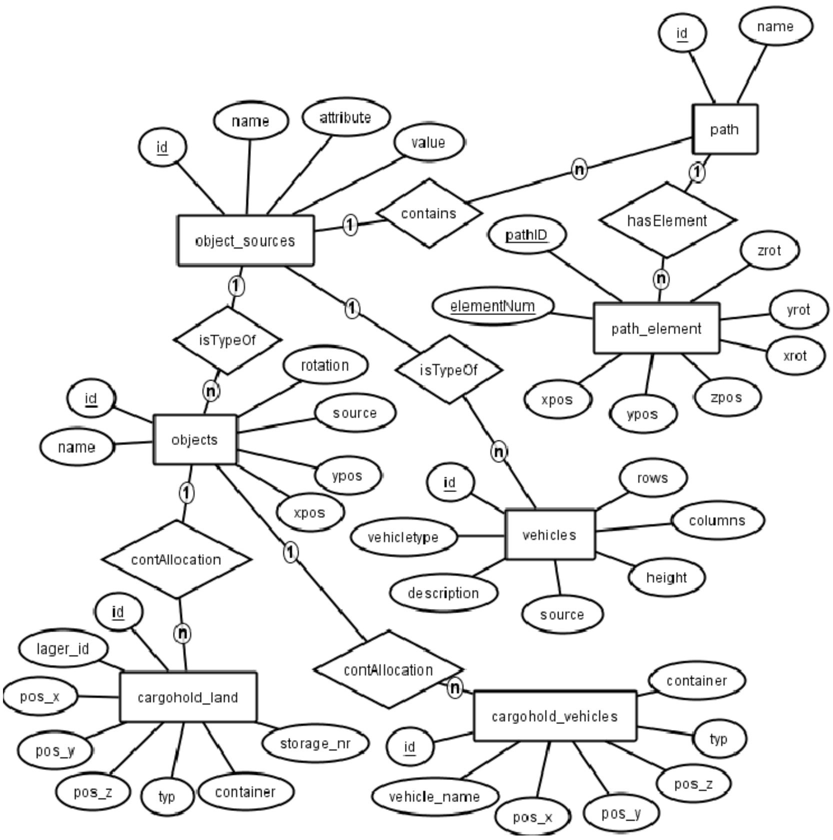


Figure 2. SIM-VIS-DB Schema as Entity Relationship Model

The hole system can be used locally or being transported for onsite review, analysis and demonstration. Due to its modularity the SIN-VIS-DB and rendering engines can also be used to visualize the container terminal in a more immersive environment (cave or a simulation center as operated by our partner university of applied Science in Elsfleth).

5. CONCLUSIONS AND OUTLOOK

With virtual port a versatile simulation environment is developed that allows the integration and joint analysis of different planning tools to build/ configure an optimal Terminal Operation System, to test and to evaluate TOS for a dedicated terminal and to evaluate and optimize Terminal infrastructure. Latter as an extension and potential verification method to dedicated analytical or simulative evaluation tools (Stahlbock and Voß 2007).

For evaluation of this approach the Jade Weser Port, Germany is simulated and the setup is used to examine minimization of retention period for ship, optimization of loading and unloading, berth allocation and resource usage. We analyzed the system on its abilities to react on unforeseen events like arrival time delays or mechanical failures of operation entities.

Using a COTS event driven simulation tool reduced implementation efforts and provide a highly capable simulation environment. With a complex terminal the system was at the edge of its capabilities but managed the model complexity. The VR visualization module is easy to configure and to set up and improves understanding and interaction of the system under examination.

6. REFERENCES

Banks, Jerry. 2005. *Discrete-event system simulation*. 4. Aufl. Upper Saddle River NJ: Pearson Prentice Hall.

Bieberstein, Norbert. 2008. *Executing SOA : a practical guide for the service-oriented architect*. Upper Saddle River NJ: IBM Press/Pearson plc.

Böse, J., D. Steenken, T. Reiners, and S. Voss. 2000. Vehicle Dispatching at Seaport Container Terminals Using Evolutionary Algorithms. In *33rd (HICSS-2000) Hawaii International Conference on System Sciences*. Bd. 2.

Maui, Hawaii.

Cantrella, Giulio Erberto, Armando Cartenì, and Stefano de Luca. 2005. A simulation model for a container terminal. In . ETC Proceedings. Strasbourg: Association for European Transport.

Franz, T., Stefan Voss, and H Rölke. 2007. Market-mechanisms for integrated container terminal management. In *Market-mechanisms for integrated container terminal management.*, 234-248. Cortona.

Günther, Hans-Otto, and Kap Hwan Kim, Hrsg. 2005. *Container Terminals and Automated Transport Systems.* Berlin/Heidelberg: Springer-Verlag. http://www.springerlink.com/index/10.1007/b137951.

Henesey, Lawrence. 2006. *Multi-agent systems for container terminal management.* Karlskrona :: Blekinge Institute of Technology,

Henesey, Lawrence, Paul Davidsson, and Jan A. Persson. 2008. Agent based simulation architecture for evaluating operational policies in transshipping containers. *Autonomous Agents and Multi-Agent Systems* 18, Nr. 2 (4): 220-238. doi:10.1007/s10458-008-9044-y.

Jammes, F., and H. Smit. 2005. Service-oriented paradigms in industrial automation. *IEEE Transactions on Industrial Informatics* 1, Nr. 1: 62–70.

Kim, K.H., and K.C. Moon. 2003. Berth scheduling by simulated annealing. *Transportation Research* 37: 541-560.

Meier, L., and R. Schumann. 2007. Coordination of Interdependent Planning Systems, a Case Study. In *INFORMATIK 2007 Informatik trifft Logistik Band 1 Beiträge der 37. Jahrestagung der Gesellschaft für Informatik e.V. (GI) 24. - 27. September 2007 in Bremen*, 389 -296. Gesellschaft für Informatik. http://subs.emis.de/LNI/Proceedings/Proceedings109/gi-proc-109-071.pdf.

O. A. o. J. CONTROLS Port Terminal Emulation Tool. http://www.tba.nl/index.php?mid=30&lang=en.

Peterkofsky, R.I., and C.F. Daganzo. 1990. A branch and bound solution method for the crane scheduling problem. *Transportation Research* 24(3): 159-172.

Ryan, Marie-Laure. 2001. *Narrative as virtual reality : immersion and interactivity in literature and electronic media.* Baltimore: Johns Hopkins University Press.

Stahlbock, Robert, and Stefan Voß. 2007. Operations research at container terminals: a literature update. *OR Spectrum* 30, Nr. 1 (10): 1-52. doi:10.1007/s00291-007-0100-9.

Zhang, C, J Liu, Y Wan, K Murty, and R Linn. 2003. Storage space allocation in container terminals. *Transportation Research Part B: Methodological* 37, Nr. 10 (12): 883-903. doi:10.1016/S0191-2615(02)00089-9.

Zhu, Y., and A. Lim. 2004. Crane scheduling with spatial constraints: mathematical model and solving approaches. In . Fort Lauderdale.

THE INTEGRATED TRUCK SEQUENCING AND STORAGE LOCATION OF EXPORT CONTAINERS PROBLEM IN A CONTAINER TERMINAL

Young-Man Park and Jong-Wook Bae

Email: ymanpark@pusan.ac.kr

Abstract: In order to improve the service for a ship owners and a trucker, the container seaport operators are trying to reduce the time of loading operation of export containers and the waiting time of container trucks. The storage location of export containers is decided so as to minimize the travel time of yard cranes in loading operation. The waiting time of trucks, which is dependent on the truck sequencing is affected by the travel time of yard crane in retrieving operation. From the yard crane's operational point of view, they are highly interdependent. Storage location of export containers has to balance the loading operation time and waiting time of container trucks. However, truck sequencing and storage location of the corresponding export container have been studied separately. Therefore, this study deals with the integrated truck sequencing and location of export containers problem in a container terminal. This paper proposed a mixed integer programming model handling truck sequencing and storage location of export containers as a whole. The objective of this model is to reduce travel time of yard crane during loading operation and waiting time of container trucks during retrieving operation. Also, we develop the heuristic algorithm for sequencing and location problem to real world adaption. A numerical example is provided to illustrate the solution procedure. We compare the heuristic algorithm with the optimum model in terms of the computation times and total weighted time. For the sensitivity analysis of configuration of storage and time weight, a variety of scenarios are experimented.

Invited Speech

·Day2: Sep. 16 (Thu.)

·Time: 16:20 - 17:40

·Chair: Il-Kyeong Moon

·Room: Ballroom C, 5F

MARITIME TRADE EVOLUTIONS AND PORT CITY DEVELOPMENTS IN ASIA

Loon Ching Tang

Department of Industrial Engineering
National University of Singapore
SINGAPORE
Email: isetlc@nus.edu.sg

Abstract: Historically, almost all goods transported worldwide have been carried by sea with the current estimate stands at approximately 90 percent by volume and 70 percent by worth. Maritime industry is an important economic sector as it has a direct impact on the prosperity of a region and/or city. We present a review on maritime trade evolution in Asia from the thirteenth centuries to the post-World War II, followed by an examination on the contemporary development of some major Asia ports. From the extant port literature, a list of factors affecting port competition and development is identified and reviewed. We then discuss future trend, challenges and opportunities facing the Asia maritime trade industry.