The Asian Journal of Technology Management Vol. 1 No. 2 (2008) 98-113

Solving Hub Network Problem Using Genetic Algorithm

Mursyid Hasan Basri*
School of Business and Management, Institut Teknologi Bandung, Indonesia

ABSTRACT

This paper addresses a network problem that described as follows. There are n ports that interact, and p of those will be designated as hubs. All hubs are fully interconnected. Each spoke will be allocated to only one of available hubs. Direct connection between two spokes is allowed only if they are allocated to the same hub. The latter is a distinct characteristic that differs it from pure hub-and-spoke system. In case of pure hub-and-spoke system, direct connection between two spokes is not allowed. The problem is where to locate hub ports and to which hub a spoke should be allocated so that total transportation cost is minimum.

In the first model, there are some additional aspects are taken into consideration in order to achieve a better representation of the problem. The first, weekly service should be accomplished. Secondly, various vessel types should be considered. The last, a concept of inter-hub discount factor is introduced. Regarding the last aspect, it represents cost reduction factor at hub ports due to economies of scale. In practice, it is common that the cost rate for inter-hub movement is less than the cost rate for movement between hub and origin/destination. In this first model, inter-hub discount factor is assumed independent with amount of flows on inter-hub links (denoted as flow-independent discount policy). The results indicated that the patterns of enlargement of container ship size, to some degree, are similar with those in Kurokawa study. However, with regard to hub locations, the results have not represented the real practice.

In the proposed model, unsatisfactory result on hub locations is addressed. One aspect that could possibly be improved to find better hub locations is inter-hub discount factor. Then inter-hub discount factor is assumed to depend on amount of inter-hub flows (denoted as flow-dependent discount policy). There are two discount functions examined in this paper. Both functions are characterized by non-linearity, so there is no guarantee to find the optimal solution. Moreover, it has generated a great number of variables. Therefore, a heuristic method is required to find near optimal solution with reasonable computation time. For this reason, a genetic algorithm (GA)-based procedure is proposed.

The proposed procedure then is applied to the same problem as discussed in the basic model. The results indicated that there is significant improvement on hub locations. Flows are successfully consolidated to several big ports as expected. With regards to spoke allocations, however, spokes are not fairly allocated.

Keywords*: Hub and Spoke Model; Marine Transportation; Genetic Algorithm*

* Corresponding author. Email: mursyid@sbm-itb.ac.id

Introduction

Nowadays there has been increasing demand for transport of passenger and goods. It takes place not only in local market, but also at regional and international market. In turn, this development of global market has generated significant economies of scale, from which any mode of transportation system can take benefits. The resulting massive flow has been reduced unit-shipping cost. This is likely occurred also due to technology development in logistic system and an increase of vehicle capacity to handle such big demands.

However, the increase of vehicle capacity requires infrastructure facilities be significantly improved and loading factor be adjusted. For an example, water depth of port is critical for large vessels. As a result, some large vessels cannot visit all ports. In addition, large vessels require efficient loading factor to reduce unit cost. Thus, the way passengers or goods transported from origin to their destination are changed. Transshipment node where consolidation activities occur is inevitably. This leads to the development of hub-and-spoke system (HSS). Vehicles with large capacity serve the links between hubs while smaller vehicles do as a feeder service to carry loads from spokes to hub.

So far, HSS is one of research topic that addressed by many investigators. In the area of air transportation, there are Hansen and Kanafani (1990), Dobson and Lederer (1993), Jaillet et al.(1996), and Bania et al.(1998), to name a few. Since the nature of the problem, HSS has also been found in the area of telecommunication (Lee et al., 1996; Klincewicz, 1998), cargo delivery (Kuby and Gray, 1993), and postal delivery service (Ernst and Krishnamoorthy, 1996). For marine transportation, several studies of HSS have been carried out by Yamato et al. (1998), Kurokawa et al. (1999).

It was reported that application of HSS lead to some benefits such as enabling shipping airlines to take advantage of economies of aircraft size (Kanafani and Ghobrial, 1985), increase in airline profitability, cost saving, increased flight frequency (quoted in Bania et al., 1998). Since marine transportation network as illustrated by Kurokawa et al. has similar characteristic in some extents, it is expected that it enjoy such benefits as well.

One factor that differs between airline and marine transport is activity at the transshipment point, where goods are charged during unloading and loading activities. As a result, HSS in marine transportation should pay higher cost for the entire system than if the system is run by direct shipping. Favorably, port authorities usually offer discounted charge if volume of flow over a certain limit. This might be a power of HSS to balance high cost at the ports. The discount policy and cost structure at each port are believed to have very important role when developing a network at the minimum total cost. Therefore, a study on hub-and-spoke network is required to develop a model that clarifies the benefits of HSS for marine transportation.

Problem Definition

Hub location problems (Hub-and-spoke system design) deal with two important decisions that are where to place the hub and how to route the flow between origins and destinations over the resulting network (O'Kelly, 1998). The former refers to location decision while the latter refers to allocation decision.

In its standard topology, hub network consists of a relatively large number of nodes, and a small number of these nodes are to be designated as hubs. Each node in the network can interact with another only via the hubs. Each non-hub node (spoke) is allocated to only one of available hubs. All hubs are fully interconnected. Then, the hub location problems generally involve finding the optimal locations for the hubs and assigning spokes to

the hubs in order to minimize the total cost of flow through the network.

There are many types of HSS. According to O'Kelly and Miller (1994), in general there are three criteria to classify hub location problems:

1. **Inter-hub connectivity**. This criterion deals with interactions among hub nodes. Based on this criterion one could find model with full or partial connectivity. In the former case any hub can be connected to all other hubs while some interaction between hubs are prohibited in the latter.

2. **Spoke-Hub allocation**. This criterion concerns with how spokes are allocated to the hubs. If each spoke is allocated only to one hub, which means each spoke can deliver and receive flow via only a single hub, it refers to single allocation model. As the opposite, in the multiple allocation model a spoke could be allocated to more than one hub, which means each spoke can deliver and receive flow via more than one hub.

3. **Inter-nodal connections**. This criterion defines interaction between two spokes. With this criterion one may allow direct connection between two spokes which bypass the hub structure. For the opposite, as applied in the standard hub-and-spoke configuration, direct connection between two spokes is not allowed.

Based on these variables, there will be eight possible systems as described in Table 1.

Since it was started by O'Kelly (1986, 1987) great deals of studies on single allocation model have been carried out by many investigators. O'Kelly (1987) formulated single allocation model with quadratic integer formulation. Two heuristic approaches were applied to solve the problem. The first heuristic, HEUR1, provides an upper bound on the objective function by allocating non-hub node to the nearest hub. The second heuristic, HEUR2, examines both the first and second nearest hub for each node with respect to the allocation part. Aykin (1990) later provided brief review on the work of O'Kelly. Klincewicz (1991) developed Exchange Heuristics and clustering approach-based heuristic to this problem. The exchange procedures (either single or double) evaluate only potential sets of hub to improve the best result found thus far. The procedure stops when no improvement could be obtained by such exchanges. Further, an attempt to obtain solution has been done by Skorin-Kapov and Skorin-Kapov (1994). They proposed a new heuristic method based on tabu search (denoted as TabuHub). It was reported that generally the results are superior to those of HEUR1 and HEUR2 in terms of solutions and computation time. Compared to those of Klincewicz (1991), although TabuHub heuristic spent longer time but it obtained superior solutions.

Table 1. Hub network classification system

Design Class	Node-hub assignment	Inter-nodal connections	Inter-hub connectivity
Protocol A	Single hub only	Not allowed	Full
Protocol B	Single hub only	Not allowed	Partial
Protocol C	Single hub only	Allowed	Full
Protocol D	Single hub only	Allowed	Partial
Protocol E	Multiple hubs allowed	Not allowed	Full
Protocol F	Multiple hubs allowed	Not allowed	Partial
Protocol G	Multiple hubs allowed	Allowed	Full
Protocol H	Multiple hubs allowed	Allowed	Partial

Skorin-Kapov et al. (1996) proposed new formulation to solve the same problem. The formulations have very tight linear programming relaxation. They approved that this approach is very effective. For the single allocation case, they are able to establish optimality of all heuristic solutions obtained via TabuHub previously. O'Kelly et al. (1996) follow-up the investigation by developing exact solutions and discuss sensitivity of this solutions to the inter-hub discount factor. For this reason, the authors employ a further reduction in the size of the problem while still maintaining the desirable integer solutions to the relaxed problem.

Another technique to solve single allocation model was proposed by Smith et al. (1996). They used modified Hopfield neural network for quadratic integer programming formulation of O'Kelly. It was reported that this technique is able to consistently obtain optimal solutions to less complex problem instances.

A solution for the special case of the problem that is single allocation model with two or three hubs is proposed by Ebery (2001). The new formulation uses fewer variables, and therefore, it is able to solve problems of twice the size that previously presented before. The other interesting case is extension to multiple allocation problems, which can be found as examples in Skorin-Kapov et al. (1996), O'Kelly et al. (1996) and Ernst and Krishnamoorthy (1998).

Variants of the standard problem are not only due to the three criteria. There exist another important factors that varying the problem such as capacity, number of hubs and flow thresholds (CSIRO, 2002). In terms of capacity, different versions of capacities can be considered, for example capacities on some of the links or on the hubs. Generally term "capacity" refers to a limit on the amount of flow being collected by hubs from the spokes. Unless explicitly indicated otherwise, a problem is assumed uncapacitated. Some works on capacitated hub location problems can be found in Aykin (1994), Ernst and Krishnamoorthy (1999), and Ebery et al. (2000).

In terms of number of hubs should be opened, the problem can be classified as *p*-hub median problem (*p*HMP) or uncapacitated hub location problem (UHLP). In the *p*-hub median problem, certain number p of the nodes are required to be opened as hubs, whereas UHLP does not prescribe the number of hubs to be opened (for an example see O'Kelly, 1992; Klincewicz, 1996; Abdinnour and Venkaratamanan, 1998). The number of hubs is usually determined by minimizing cost with either capacity constraints or fixed costs for opening a hub.

In regard with thresholds, this has not been studied extensively. It requires minimum flows across some or all of the links. For example in the airline application where multiple allocation is common, flow thresholds corresponding to the smallest plane size operated by the company could be imposed to prevent uneconomical links from being included in the network (CSIRO, 2002).

Another classification is in terms of the type of space in which hubs are located (Sasaki and Fukushima, 2001). Based on this criterion, a problem could be classified as discrete or continuous location problem. In the former, a hub can be located only at one of a finite number of candidate nodes, while in the latter a hub can be located arbitrarily in a region on a plane. For examples of discrete hub location problems, Campbell (1994) mentioned four types: *p*-hub median problem, uncapacitated location problem, *p*-hub center problems and hub covering problems.

Last, the overall classification of location models is proposed by Hamacher and Nickel (1998). They proposed 5-position classification scheme so that not only classes of specific location models are covered, but also all of them in a single scheme.

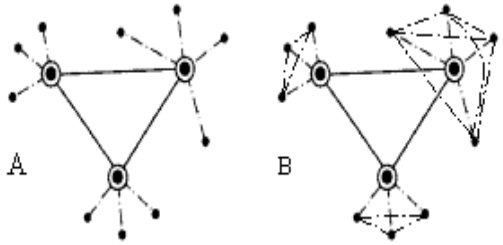
This paper presents hub network design for marine transportation. Therefore, before selecting one best suit model for the problem, it is necessary to look at marine transportation network model. One of reasonable models is network model of Kurokawa et al. (1999). It is one of discrete *p*-hub median problem. Refer to Table 1., it has the characteristic of single allocation model and fully interconnected hubs. In regards with inter-nodal connections, it allows inter-nodal connection. However, such connections are limited only for nodes (ports) that allocated to the same hub. As a result, the problem encountered here is a variant of single allocation models in Table 1.

Unlike most of works on *p*-hub location model, Kurokawa et al. determined hub locations by judgment of designer after doing cluster analysis based on distance between each pair of origin-destination. This implied that process of selecting hub location ignores amount of flows throughout the network. Aykin (1995) indicated that such approach is not necessarily to obtain an optimal solution. Inspired to find optimal hub location and solve the problem simultaneously, a model for marine transportation network is needed.

Mathematical Formulation

Basic Model

The main reasons for developing basic model are to provide an optimal model for the problem and a basis for comparison with Kurokawa model. As mentioned before, it differs from those in hub location studies mainly in terms of its structure. Kurokawa et al. allowed direct connections between two non-hub nodes that allocated to the same hub (Fig. 1.B), while previous investigators permitted no such connections (Fig. 1.A).



A. Pure Hub & Spoke B. Proposed model

Figure 1. Network configurations

The basic model was developed by combining main features of Kurokawa and Skorin-Kapov (1996) models. Figure 2 shows how basic model was developed based on both models.

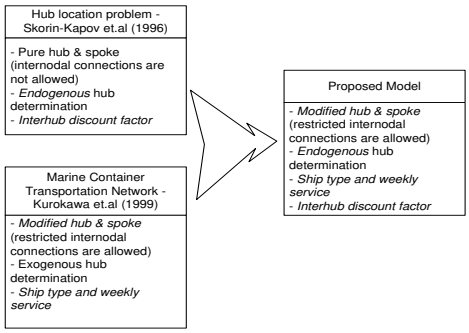


Figure 2. Research position

Objective Function

The objective function is to minimize total transportation cost. Since it was based on Skorin-Kapov model, modification from original function is required due to permission of restricted inter-nodal connection and consideration of empty container flow to balance the transportation system. The new total cost is formulated as follow.

$$Minimize\ TC = C1 + C2 + C3 \qquad (1.1)$$

Where:

$$C1 = \sum_i \sum_k \sum_{l|l\neq k} \sum_j \sum_s Y_{ikljs}(C_{ik} + \alpha\, C_{kl} + C_{lj}) \qquad (1.2)$$

$$C2 = \sum_i \sum_k \sum_{l|l=k} \sum_j \sum_s Y_{ikljs} C_{ij} \qquad (1.3)$$

$$C3 = \sum_i \sum_j \sum_s X_{ijs} C_{ij} \qquad (1.4)$$

TC : Total cost [yen/year]
C1 : Inter-hub cost [yen/year]
C2 : Inter-nodal cost [yen/year]
C3 : Empty container cost [yen/year]
Y_{ikljs} : *the flow from node i to node j via hubs k and l using ship type s [TEU/year]*
X_{ijs} : *the flow of empty container from node i to node j using ship type s [TEU/year]*
C_{ij} : *cost per unit of flow for link ij*
α : *Coefficient of inter-hub discount factor*

Each cost component contains 5 elements: fuel, harbor, handling, ship, and container cost. Index *s* is added to express ship type that should be assigned for each link.

A Set of Constraints

Set of constraints of the basic model is expressed in equation (1.5)-(1.14).

$$\sum_k H_{kk} = p, \qquad (1.5)$$

$$\sum_k H_{ik} = 1, \forall i \in N \qquad (1.6)$$

$$H_{ik} \le H_{kk}, \forall i,k \in N \qquad (1.7)$$

$$\sum_l \sum_s Y_{ikljs} = w_{ij} H_{ik}, \forall i,k,j \in N \qquad (1.8)$$

$$\sum_k \sum_s Y_{iklj} = w_{ij} H_{jl}, \forall i,l,j \in N \qquad (1.9)$$

$$\sum_j \sum_s X_{ijs} = SEC_i, \forall i \in N \qquad (1.10)$$

$$\sum_i \sum_s X_{ijs} = DEC_j, \forall j \in N \qquad (1.11)$$

Where:

$$Y_{ikljs}, X_{ijs} = 0 \text{ for } h_i \le d_s \qquad (1.12)$$

$$H_{ik} = \begin{Bmatrix} 1 \text{ if node } i \text{ is allocated to hub } k, \\ 0 \text{ otherwise} \end{Bmatrix}, \qquad (1.13)$$

$$Y_{ikljs}, X_{ijs} \ge 0, \forall i,j,k,l \in N; \forall s \in S \qquad (1.14)$$

N : number of nodes
p : number of hubs
w_{ij} : *the flow from i to j*
h_i : *depth of water of port i*
d_s : *draft of ship size s*

Eq. (1.5) specifies the number of hubs. Eq. (1.6) is the single allocation constraint. Eq. (1.7) ensures that a node may only be allocated to a hub. Eqs. (1.8) and (1.9) ensure that flow between two nodes is delivered. Eqs. (1.10) and (1.11) ensure that empty container distribution satisfies demand and supply sides. Eq. (1.12) represents water depth restriction. By assuming that demanded empty container node can be supplied from any node, empty container flow will become common linear programming model for transportation problem (Winston, 1994). For this purpose the following formulation are applied:

$$SEC_i = \begin{Bmatrix} D_i - O_i, if\ D_i - O_i \ge 0 \\ 0\ otherwise \end{Bmatrix}, \qquad (1.15)$$

$$DEC_i = \begin{Bmatrix} 0\ if\ D_i - O_i \ge 0 \\ O_i - D_i\ otherwise \end{Bmatrix}, \qquad (1.16)$$

Where:

SEC_i : *surplus of empty container at node i*
DEC_i : *demand of empty container at node i*

$$O_i = \sum_j w_{ij}, \forall i \in N \qquad (1.17)$$

$$D_i = \sum_j w_{ji}, \forall i \in N \qquad (1.18)$$

Flow-Dependent Discount Model

The problem formulated in this section is similar to basic model. This work differs from the basic version by application of flow-dependent discount function α. In the present work α is only applied to harbor cost and container handling cost as representation of port cost. This condition is more relevant to some real cases that discounted fee is charged for port-related costs.

To avoid larger ship selected to serve branch line, ship size is dropped as index of decision variable. Accordingly, ship size will be assigned separately: the smaller for feeder service and the larger for inter-hub service.

Since empty container movement was assumed can be handled by direct shipping, its optimization can be carried out independently. For that reason, in this work optimization of empty container is excluded.

With these additional considerations, the problem is stated as follows:

$$Minimize\ TC = C_1 + C_2 \qquad (2.1)$$

C_1 (Cost between two ports via hubs) and C_2 (Cost between two non-hub ports that allocated to the same hub) are defined by

$$C_1 = \sum_i \sum_k \sum_{l|l\neq k} \sum_j X_{iklj}(f_{fuel} + f_{harb}(\alpha) + f_{hand}(\alpha) + f_{ship} + f_{cont}) \qquad (2.2)$$

$$C_2 = \sum_i \sum_k \sum_{l|l=k} \sum_j X_{iklj}(f_{fuel} + f_{harb} + f_{hand} + f_{ship} + f_{cont}) \qquad (2.3)$$

Where:

X_{iklj} : *the flow from node i to node j via hubs k and l*

α : *coefficient of inter-hub discount factor*

Functions f_{fuel}, f_{harb}, f_{hand}, f_{ship} and f_{cont} represent cost component of fuel cost, harbor cost, container handling cost, ship cost, and container cost respectively. To avoid improper calculation for α-related cost, some restrictions should be added:

Table 2.a Detailed calculation for harbor cost

Conditions		Calculation
i=k	i=j	αf_{harb-l}
i=k	l≠j	$\alpha f_{harb-l} + f_{harb-j}$
i≠k	l=j	$\alpha f_{harb-k} + \alpha f_{harb-l}$
i≠k	l≠j	$\alpha f_{harb-k} + \alpha f_{harb-l} + f_{harb-j}$

Table 2.b Detailed calculation for container handling cost

Conditions		Calculation
i=k	l=j	$\alpha f_{hand-k} + \alpha f_{hand-l}$
i=k	l≠j	$\alpha f_{hand-k} + 2\alpha f_{hand-l} + f_{hand-j}$
i≠k	l=j	$f_{hand-i} + 2\alpha f_{hand-k} + \alpha f_{hand-l}$
i≠k	l≠j	$f_{hand-i} + 2\alpha f_{hand-k} + 2\alpha f_{hand-l} + f_{hand-j}$

Sets of constraints for the problem are follows:

$$\sum_k H_{kk} = p, \qquad (2.4)$$

$$\sum_k H_{ik} = 1, \forall i \in N \qquad (2.5)$$

$$H_{ik} \leq H_{kk}, \forall i,k \in N \qquad (2.6)$$

$$\sum_l X_{iklj} = w_{ij} H_{ik}, \forall i,k,j \in N \qquad (2.7)$$

$$\sum_k X_{iklj} = w_{ij} H_{jl}, \forall i,l,j \in N \qquad (2.8)$$

$$X_{iklj} = 0 \text{ for } h_i \le d \quad (2.9)$$

$$H_{ik} = \begin{cases} 1 \text{ if node } i \text{ is allocated to hub } k, \\ 0 \text{ otherwise} \end{cases} \quad (2.10)$$

$$X_{iklj} \ge 0, \forall i, j, k, l \in N \quad (2.11)$$

where:

N : *number of nodes*
p : *number of hubs*
w_{ij} : *the flow from i to j*
h_i : *depth of water of port i*
d : *draft of ship*

All parameters related to each component of cost function are similar with those in the basic model.

Inter-hub Discount Function

According to O'Kelly and Bryan flow-dependent discount will be better in locating the hubs and estimating total network cost. They proposed the discount rate (ϕ) in non-linear form expressed as follow:

$$\phi = \theta \left[\frac{X_{ij}}{\sum_i \sum_j X_{ij}} \right]^{\beta} \quad (2.12)$$

Where

ϕ *=1-*α *(*α *inter-hub discount factor)(2.13)*

θ *and* β *Parameters;* $\theta > 0$, $\beta > 0$

X_{ij} *An amount of flow between origin i and destination j*

This function will give an increasing discount at decreasing rate. By tuning value of parameter θ and β, appropriate discount can be determined.

Another formulation to express discount function is extracted from some information available in some ports website (Port of Yokohama, Port of Salalah, Port Louis – Republic of Mauritius). Instead of using non-linear function, discount function empirically results in step function. For an example, port dues will be discount 50% for vessels handling more than 1,500 TEU, and discount 30% for vessels handling between 1,000 and 1,499 TEU (Port of Yokohama). The maximum discount given at Port of Yokohama and Port of Salalah are 50%, while Port Louis offer maximum 40% discount (with possibility to get higher discount for large transshipment exchange). Based on this information step function for inter-hub discount rate is set as follow:

$$\phi = \begin{cases} 0 & X_{ij} < 182{,}500 TEU / year \\ 0.1 & 182{,}500 \le X_{ij} < 365{,}000 TEU / year \\ 0.3 & 365{,}000 \le X_{ij} < 547{,}500 TEU / year \\ 0.5 & X_{ij} \ge 547{,}500 TEU / year \end{cases} \quad (2.14)$$

GA Formulation

Given the flow and distance table between each pair of the network understudy with n ports, the problem is to select location of p hubs and allocate the remaining (n-p) ports to a single hub at the minimum cost. The main parts of GA model proposed in this work are as follows:

Representation.

To represent an individual as a chromosome, a series of integer number (p+m) are generated. Each chromosome contains two parts: the first part that consists of p digits (*genes*) express position of hubs, and the second part with m genes express a network configuration, in which allocation of non-hub ports are defined. Fig. 3. illustrates a network with 6 ports and 3 hubs. Locations of hubs are determined by first 3 genes of chromosome: 1st node, 2nd node and 4th node that are port 4, 2 and 3 respectively. The remaining non-hub ports are allocated with the following rule: any non-hub port is allocated to the hubs at the left

relative to its position. In this example, port 6 is allocated to hub-port 2 and port 1 and 5 are allocated to hub-port 3.

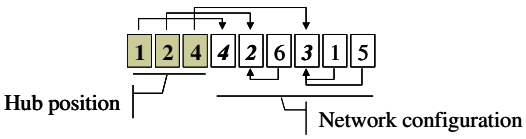


Figure 3. Example of chromosome representation

Fitness function

Total cost of the system is clearly defined by $\sum_i \sum_j X_{ij} C_{ij}$. Since the problem is minimization, modification of fitness function is required. Here fitness function is defined by

$$ff = BIG - \sum_i \sum_j X_{ij} C_{ij} \qquad (2.15)$$

where BIG is a positive number that larger than $\sum_i \sum_j X_{ij} C_{ij}$.

Variable X_{ij} is an amount of bundled flows from i to j and resulted from modification of OD matrix according to a particular network. Therefore, fitness function can be easily calculated for any network configuration, which has unique equivalent chromosome.

Genetic operators

Three main genetic operators as underlying fundamental mechanism of GAs i.e. selection, crossover and mutation are employed. In this work binary tournament selection (tournaments are held between two individuals) is applied in selection process (Blickle and Thiele, 1995). Two individuals are randomly chosen from the population. Individual having higher fitness value is chosen deterministically and inserted into the next population. The process is repeated as much as required to obtain a new population.

Since integer value is used in chromosome, classical crossover does not work. Therefore another technique to implement crossover is required. One of the most popular is the PMX - Partially Matched Crossover (Goldberg and Lingle, 1985). First, as other methods, two chromosomes are aligned. Two crossing points are selected at random along the second part of chromosome to define matching area. Then the genes in this area are exchange. Fig. 4(a) shows output of all these process. Finally, to get a valid permutation a repair should be done as illustrated in Fig. 4(b).

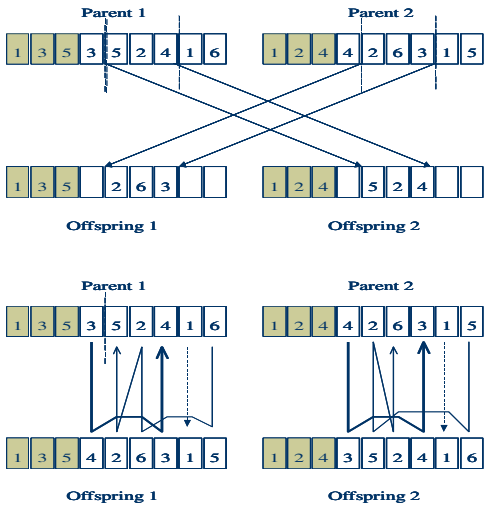


Figure 4. Partially Matched Crossover

For the last operator, simple inversion mutation (Larranaga et al., 1999) is used in this study. The mechanism starts with selecting two cut points at random. Then, genes located between two cut points are reversed (See Fig. 5.).

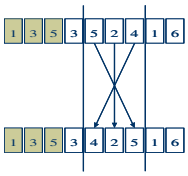


Figure 5. Simple Inversion Mutation

The procedure of GA in the current work is described as followed. It starts with inputting GA parameters: number of iteration ($Iter_{max}$), number of individuals in each generation (Pop_{max}), number of generation (Gen_{max}), crossover and mutation probabilities. Number of iteration is required when GA procedure will be run repeatedly for different random seed number. Pop_{max} is set 40 in this work.

Initialization generates a random initial population. It also includes evaluation on fitness of each individual. The solutions in initial population provide a baseline to judge future search efforts. On each generation, crossover and mutation mechanisms are applied to individuals, which are probabilistically selected from the population. Then, fitness of each new individual is evaluated. Over one population, statistic is collected and recorded. The best fitness so far is also kept until the last generation. The procedure stops when number of iteration is achieved.

Data Set and Scenario

Data set for this study includes 19 ports with high traffic in 1994 (Kurokawa et al., 1999). Most of the ports are in the East and Southeast Asian area. Several ports outside this area e.g. North American and European area are included because intensive trade exists among them. Fig. 6 shows location of the ports.

Scenarios developed for second model consist of 2 categories. The first, called scenario A, is aimed to verify if GA model works well to solve the problem. This only possible when the result obtained is compared with that of LP-based optimal model. Therefore, parameter α is set to 1. Since scenario A is unnecessarily to consider discount factor, the same ship size (1500 TEU) for feeder service and inter-hub service is employed.

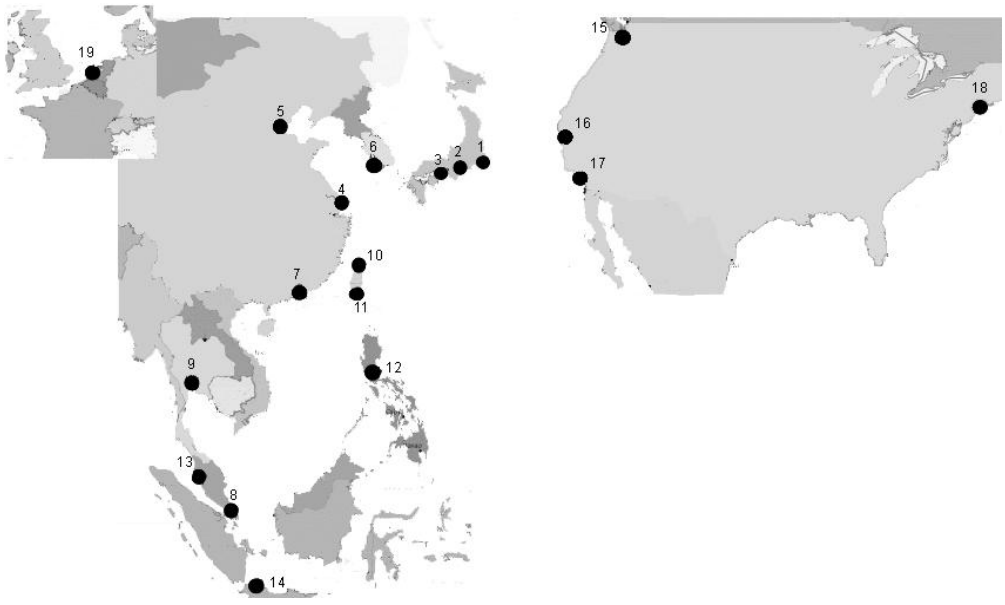


Figure 6. List of Ports:Tokyo/Yokohama (2) Nagoya (3) Osaka/Kobe (4) Shanghai (5) Tianjin (6) Busan (7) Hong Kong (8) Singapore (9) Bangkok (10) Keelung (11) Kaohsiung (12) Manila (13) Port Klang (14) Tanjung Priok (15) Seattle/Tacoma (16) Oakland (17) Los Angeles/Long Beach (18) New York (19) Rotterdam

The second, called scenario B, is aimed to evaluate impact of flow-dependent discount function for the system. For this purpose, two types of non-linear function are evaluated. Scenario with function expressed in (2.12) is called B1 while scenario with function expressed in (2.14) is called B2. From various possible values of θ and β in scenario B1, two functions are evaluated (θ =1; β=0.3 and β=0.4). Unlike scenario A, 1500 TEU ship is used to serve feeder service and 3000 TEU ship to serve inter-hub service. We assumed the same step function is applied for harbor cost and container handling cost.

Model Verification

To verify the model, results of scenario A are compared with those obtained with optimal model that solved by software package LINGO 8. All cases used in this verification belong to flow-independent discount cost model with α=1 and presented in Table 3. The table shows that proposed GA model is able to obtain optimal solution in all cases. Although for small size of the problem it takes longer time, as problem size bigger its computation time much better than optimal model. Therefore, it is concluded that GA performance is acceptable to solve the problem.

Discussion

Basic Model

From the results presented in Fig. 7(a)-(b), it can be seen that total costs of proposed model are lower than those in the previous study. This is not surprising because predetermined hub location as conducted by Kurokawa et al. cannot guarantee optimality of the solution, besides the procedure requires very intensive investigation to select the one best hub in each group of ports. Those figures, however, demonstrated that the proposed model could be a promising alternative method to obtain a better network configuration.

The result of application of water depth restriction (Δ) indicated that there is similar pattern between both models. For container ships smaller than 5,000 TEU the total cost tends to decrease and then start to increase as larger ship is operated. As seen in Fig. 7(a) compared to 7(b), the total costs obtained are lower than those of the previous study. This situation may occur due to some reasons. First, locations of hubs obtained by both models are different. Secondly, the proposed model searches for the size of ship that will minimize total cost either in main line or branch line. In the previous study main line and branch line are served by predetermined ship size. It is shown that either inter-hub flow or direct flow between ports use a larger ship as long as water depth of port is not violated.

Table 3. GA performance compared with optimal model

α	n	p	Solution (10^3)		Computation time (sec)	
			Optimal	GA	Optimal	GA
1	5	2	3,049	3,049	3	3.5
		3	3,452	3,452	4	5.1
	10	3	19,521	19,521	106	7.212
		4	20,384	20,384	285	14.752
		5	21,278	21,278	280	5.569
		6	21,129	21,129	621	7.252
	19	3	76,243	76,241	4 hrs 15 min	30.505
		4	77,054	77,053	6 hrs 51 min	17.206
		5	77,909	77,908	6 hrs 42 min	17.406
		6	78,880	78,879	10 hrs 13 min	16.174

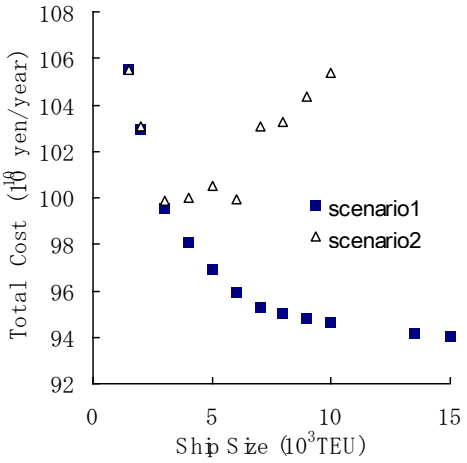
Increasing trend of total cost for container ships bigger than 5,000 TEU in both models is caused by an application of water depth restriction. In most cases, due to larger ship cannot enter port, 1,500 TEU ship is operated to carry the containers. This leads us to the previous case where the operation cost of smaller ship is higher. For an example, let see the case of 1,500 TEU and 10,000 TEU. The total cost is not far from base-case where only use 1,500 TEU ship. With water depth restriction, 10,000 TEU ship can only enter Port Los Angeles and Rotterdam. This implied that most links are served by 1,500 TEU ship.

The results show total transportation cost of proposed model was lower than those of the previous study for both scenarios. It seems dominance of container handling cost impel the system to select ports that have cheap handling cost and serve relatively small quantity of containers. This situation also let many spokes be allocated to one hub port in order to enjoy direct connection where loading/unloading activities at hub port are not necessary. Flow-Dependent Discount Model

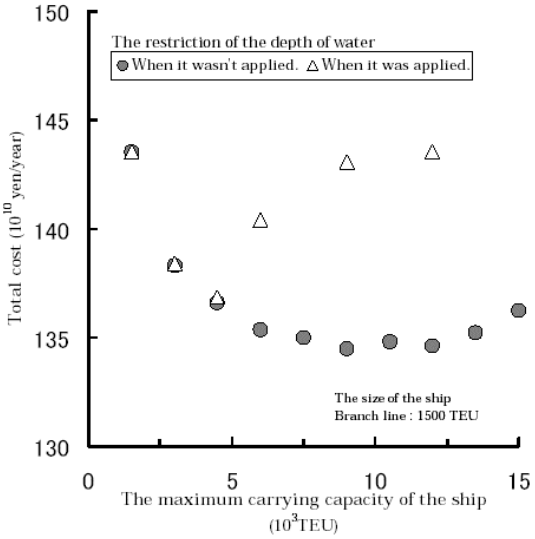
Trend of Total Cost

In this section, trend of total cost versus number of hubs is presented. This trend is important to see if the advantage of HSS occurs. Ideally, the performance of a HSS outperforms that of a system without hubs. This implied that total cost should decrease for certain range of number of hubs.

Fig. 8 shows the trend of total cost for two scenarios as number of hubs increase. Both of Scenario B1 demonstrates decreasing total cost as p increases while scenario B2 indicates a different nature. The latter indicated that total cost reach minimum value at p=6. Trend of Scenario B1 seems to be unrealistic rather than that of B2. Although at the best our knowledge there is no empirical or theoretical study related to this, it sounds benefits of HSS is only applicable for a limited range of p.



a. Proposed model



b. Kurokawa et al. result (English translation)

Figure 7. The examination on the enlargement of the container ship

The contrast trend showed between these two scenarios is likely due to different characteristic of underlying function. Scenario B1 always enjoys discount regardless amount of flow transported between hubs. In addition, in general increasing hubs mean more links enjoy benefits of direct shipping. Meanwhile,

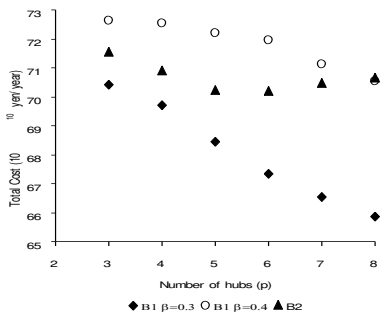


Figure 8. Trend of total cost when number of hubs increase

scenario B2 receives discounted charge only if amount of flow over a certain limit.

Table 4 shows components of total cost and gap of two most dominant cost components. The gap measures change of cost component when additional hub is considered. This table emphasizes hypothesis described earlier. Almost all cases in scenario B1 indicate that system enjoys discount as shown by decreasing of container handling cost. This table also shows that scenario B2 only enjoys discount until p=5. Although ship cost tends to decrease, the dominance of container handling cost outweigh it. As a result, total cost start to increase from p=6.

Hub location decision

As the main objective, decision on hub location is quite important. In the basic model, it failed to select big hub ports. It is expected that under flow-dependent discount policy, proposed model obtain specified results.

Due to some change in the nature of model, a direct quantitative comparison with previous model would be meaningless. However, comparison between results on hub location with the previous is useful to see if proposed model obtain better result. For that reason, only results when p=6 is depicted and presented in Table 5.

Table 4. Cost components for all scenarios

Case No	Scenario	p	Cost (10^{10} yen/year)						Gap*	
			Total	Fuel	Harbor Cont. Handl.		Ship	Container Cont. Handl.		Ship
1		*3*	704.28	29.71	8.95	494.48	135.55	35.60		
2		*4*	697.14	32.24	8.30	486.60	133.58	36.42	(7.88)	(1.97)
3	B1 β=0.3	*5*	684.52	32.26	8.20	474.74	132.92	36.40	(11.86)	(0.65)
4		*6*	673.35	32.76	7.76	462.53	133.51	36.78	(12.20)	0.59
5		*7*	665.37	32.78	7.65	461.32	127.80	35.83	(1.22)	(5.71)
6		*8*	658.75	32.77	7.59	459.50	123.76	35.14	(1.82)	(4.04)
7		*3*	726.43	30.14	9.04	520.97	131.36	34.92		
8		*4*	725.39	30.22	9.21	518.83	131.91	35.22	(2.14)	0.55
9	B1 β=0.4	*5*	722.11	32.26	8.86	511.67	132.92	36.40	(7.16)	1.02
10		*6*	719.61	32.71	8.71	513.26	128.99	35.94	1.59	(3.94)
11		*7*	711.15	32.78	8.43	506.32	127.80	35.83	(6.94)	(1.19)
12		*8*	705.45	32.77	8.37	505.41	123.76	35.14	(0.91)	(4.04)
13		*3*	715.60	29.71	9.13	505.62	135.55	35.60		
14		*4*	709.11	32.24	8.56	498.32	133.58	36.42	(7.30)	(1.97)
15	B2	*5*	702.33	32.26	8.52	492.22	132.92	36.40	(6.10)	(0.65)
16		*6*	701.91	32.71	8.43	495.85	128.99	35.94	3.63	(3.94)
17		*7*	704.85	33.15	8.24	497.93	129.29	36.24	2.08	0.30
18		*8*	706.53	33.14	8.24	504.35	125.25	35.56	6.42	(4.04)

* Gap = $cost_p - cost_{p-1}$

Number in parenthesis indicates negative amounts.

Table 5. Selected hub ports

Model	Selected Hubs
Basic Model	Kaohsiung, Manila, New York, Port Klang, Tianjin, and Tanjung Priok
Scenario B1 (β=0.3)	Kaohsiung, Rotterdam, Los Angeles, Hong Kong, Tokyo, and Kobe
Scenario B1 (β=0.4)	Kaohsiung, Rotterdam, Los Angeles, Hong Kong, Tokyo, and New York
Scenario B2	Kaohsiung, Rotterdam, Los Angeles, Hong Kong, Tokyo, and New York

From Table 5 some important notes can be pointed out as follows:

- All three scenarios in the second model selected big ports as hubs. This proved that flow-dependent discount enables model to reflect real conditions better than previous one.
- All three scenarios select Hong Kong. This is important result because Hong Kong is one of the major hubs in the world.
- In terms of dispersion, obtained results are much better than those in previous models. At least each zone outside of Asia is represented by one hub port: Los Angeles for America and Rotterdam for Europe.

However, results of better hub location are not followed by fairly spoke allocations. All non-hub ports are allocated to Kaohsiung. This unsatisfactory result may occur mainly due to unlimited port capacity assumption.

Conclusion

In this paper a new discount model to gain economies of scale for marine transportation network is presented. Non-linear and step function that represent flow-dependent discount model is proposed to consolidate more flow from and to hub ports. As expected, both discount functions succeeded to bundle flows into big hub ports. This result confirmed that discounted price at port is an important factor to realize economies of scale for marine transportation network together with application of larger ship. Some encouraging findings concern with hub location decision are (i) Hong Kong as one of the biggest hub ports in the world is always selected, (ii) selected hubs are fairly distributed within area understudy such as Rotterdam for Europe and Los Angeles for America besides hub port in Asia. These findings demonstrated the appropriateness of proposed model to select hub ports.

In regards with discount function, both functions evaluated in this work resulted in similar hub locations. However, non-linear function as proposed by O'Kelly and Bryan has not represented feature of hub-and-spoke system for marine transportation adequately. In contrast, step function successfully outlined such feature. In addition, it is easier to implement practically.

As discussed earlier, concentrated flow into one central hub is occurred. Undoubtedly, this is not satisfying result in terms of spoke allocation. Therefore, in the near future we would like to investigate more comprehensive and realistic model that include other constraints such as number of berth and its capacity.

References

Abdinnour-Helm, S. and Venkataramanan, M.A. (1998). Solution Approaches to Hub Location Problems. *Annals of Operations Research*, Vol. 78, No. 0, pp. 31-50.

Aykin, T. (1990), On A Quadratic Integer Program for The Location of Interacting Hub Facilities. *European Journal of Operational Research*, Vol. 46, No. 2, pp. 409-411.

Aykin, T. (1994). Lagrangian Relaxation Based Approaches to Capacitated Hub-and-

Spoke Network Design Problem. *European Journal of Operational Research*, Vol. 79, No. 3, pp. 501-523.

Aykin, T. (1995). Networking policies for hub-and-spoke systems with application to the air transportation system. *Transportation Science*, Vol. 29 No. 3, pp. 201-221.

Bania, N., Bauer, P.W., and Zlatoper, T.J. (1998). U.S. Air Passenger Service: A taxonomy of route network, hub location and competition. *Transportation Research*, Vol. 34, No. 1, pp. 53-74.

Blickle, T., and Thiele, L. (1995). A mathematical analysis of tournament selection. *Proceedings of the Sixth International Conference on Genetic Algorithms*.

Campbell, J.F. (1994). Integer Programming Formulations of Discrete Hub Location Problems. *European Journal of Operational Research*, Vol. 72, No. 2, pp. 387-405.

CSIRO.
http://www.cmis.csiro.au/or/hubLocation/.

Dobson, G., and Lederer, P.J. (1993). Airline scheduling and routing in a hub-and-spoke system. *Transportation Science*, Vol. 27, No. 3, pp. 281-297.

Ebery, J. (2001). Solving large single allocation p-Hub problems with two or three hubs. *European Journal of Operational Research*, Vol. 128, No. 2, pp. 447-458.

Ebery, J., Krishnamoorthy, M., Ernst, A., and Boland, N. (2000). The Capacitated Multiple Allocation Hub Location Problem: Formulations and Algorithms. *European Journal of Operational Research*, 120, 614-631.

Ernst, A.T., and Krishnamoorthy, M. (1996). Efficient Algorithms for The Uncapacitated Single Allocation p-Hub Median Problem. *Location Science*, 4, pp. 139-154.

Ernst, A.T., and Krishnamoorthy, M. (1998). Exact and heuristic algorithms for the uncapacitated multiple allocation p-Hub median problem. *European Journal of Operational Research*, 104, pp. 100-112.

Ernst, A.T. and Krishnamoorthy, M. (1999). Solution algorithms for the capacitated single allocation hub location problem. *Annals of Operations Research*, 86, pp. 141-159.

Goldberg, D.E. and Lingle, R. (1985). Alleles, loci and traveling salesman problem. *Proceeding International Conference on Genetic Algorithms and Their Applications*.

Hamacher, H.W., and Nickel, S. (1998). Classification of Location Models. *Location Science*, 6, pp. 229-242.

Hansen, M. and Kanafani, A. (1990). Airline hubbing and airport economics in the Pacific market. *Transportation Research A*, Vol. 24, No. 3, pp. 217-230.

Kanafani, A. and Ghobrial, A.A. (1985). Airline hubbing - some implications for airport economics. *Transportation Research A*, Vol.19, pp. 15-27.

Klincewicz, J.G. (1991). Heuristics for the p-Hub location problem. *European Journal of Operational Research*, 53, pp. 25-37.

Klincewicz, J.G. (1998). Hub location in backbone/tributary network design: A review. *Location Science*, 6, pp. 307-335.

Klincewicz, J.G. (1996). A dual algorithm for the uncapacitated hub location problem. *Location Science*, Vol. 4, No. 3, pp. 173-184.

Kuby, M.J., and Gray, R.G. (1993). The hub network design problem with stopovers and feeders: The case of Federal Express. *Transportation Research A*, 27, pp. 1-12.

Kurokawa, H., Tsuruta, S., and Shima, K. (1999). A study on the design of the marine

transportation network - around the East Asia and the South East Asia. *Journal of the Navigation Society of Japan*, 101, pp. 259-269 (in Japanese).

Larranaga, P, Kuijpers, C.M.H., Murga, R.H., Inza, I., and Dizdarevic, S. (1999). Genetic algorithms for the travelling salesman problem: A review of representations and operators. *Artificial Intelligence Review*.

Lee, Y., Lim, B.H., and Park, J.S. (1996). A Hub location problem in designing digital data service network: Lagrangian Relaxation approach. *Location Science*, 4, pp. 185-194.

O'Kelly, M.E. (1986). Activity levels at hub facilities in interacting networks. *Geographical Analysis*, 18, pp. 343-356.

O'Kelly, M.E. (1987). A quadratic integer program for the location of interacting hub facilities. *European Journal of Operational Research*, 32, pp. 393-404.

O'Kelly, M.E. (1992). Hub facility location with fixed costs. *The Journal of The RSAI*, 71, pp. 293-306.

O'Kelly, M.E. (1998). A geographer's analysis of hub-and-spoke networks. *Journal of Transport Geography*, Vol. 6, No. 3, pp. 171-186.

O'Kelly, M.E., and Miller, H.J. (1994). The hub network design problem: A review and synthesis. Journal of Transport Geography 2(1), pp. 31-40.

O'Kelly, M.E., Bryan, D., Skorin-Kapov, D. and Skorin-Kapov, J. (1996). Hub network design with single and multiple allocation: A computational study. *Location Science*, 4, pp. 125-138.

P Jaillet, G Song, G Yu. (1996), Airline network design and hub location problems, *Location Science*, 4, pp. 195-212.

Sasaki, M. and Fukushima, M. (2001). Stackelberg hub location problem. *Journal of Operations Research*, Vol. 44, No. 4, pp. 390-401.

Skorin-Kapov, D. and Skorin-Kapov, J. (1994). On Tabu search for the location of interacting hub facilities. *European Journal of Operational Research*, 73, pp. 502-509.

Skorin-Kapov, D., Skorin-Kapov, J., and O'Kelly, M.E. (1996). Tight linear programming relaxations of uncapacitated p-Hub media problems. *European Journal of Operational Research*, 94, pp. 582-593.

Smith, K., Krishnamoorthy, M., and Palaniswami, M. (1996). Neural versus traditional approaches to the location of interacting hub facilities. *Location Science*, Vol. 4, No. 3, pp. 155-171.

Yamato, H., Tsunoda, T., Koyama, T., and Fushimi, A. (1998). A design of the container cargo flow network based on the mathematical programming. *Journal of the Society of Naval Architect of Japan*, 184, pp. 607-615 (in Japanese).