Scientiae Mathematicae Japonicae Online, **e-2006**, 475–488

A DUAL BASED HEURISTIC FOR A HUB-SPOKE AIR CARGO NETWORK DESIGN PROBLEM WITH HOP-COUNT CONSTRAINTS

Moon-Gil Yoon*, Duk-Young Yoon† and Sang-Heon Han‡

Received March 24, 2006

Abstract. In this paper, we address a hub-spoke network design problem for air-cargo systems. To build such a network, three kinds of network costs should be considered: fixed costs for establishing a hub, fixed costs for operating air-cargo on each route and variable costs occurring on each route. With these kinds of costs, we develop an optimization model for designing a hub-spoke network in air-cargo systems, including the hop-count constraint being used effectively to deliver freights. We suggest a dual based heuristic algorithm to solve our problem. Computational experiments show that the proposed heuristic is satisfactory in both speed and the quality of the solutions generated.

1 Introduction The hub-and-spoke (HS) structure has been extensively adopted in the airline industry for last two decades, and this structure has proven to be flexible and cost-effective as is evidenced by their increased use in the transportation industry. In an HS network, each traveler has a more frequent travel schedule to choose from, but it takes a longer distance and a longer time because non-stop service is reduced (Bryan and OfKelly, 1999, Sasaki et al., 1999).

For air-cargo systems, HS structure has the same benefit as for passenger airlines. An example of an air-cargo system with HS structure is given in Figure 1, where user and hub nodes correspond to local and hub airports respectively; each arc represents a flight route. The demand at a user node is transported to a hub where they are sorted and rerouted to their respective destinations. Usually, a small or a medium-sized cargo plane is assigned to transport the demand between a hub and a user node, while a large-capacity carrier is used on the route between hubs. Occasionally, a cargo plane may be used to deliver the demand directly between the origin and the destination nodes.

There are several researches on HS network design problems. One of the first papers on an HS network design problem was that of O'Kelly (1986). He dealt with the optimal locations of a single hub and two hubs in a network by minimizing flow-weighted distance. Further researches in this field include Aykin (1995), Horner and OfKelly (2001), Jaillet et al. (1996), Mayer and Wagner (2002), and Sasaki et al. (1999).

Aykin (1995) investigates two different variants of the hub design problem. In the first variant, all traffic from a given point must flow through a specific hub before proceeding to its destination. The second one permits trips from a given origin to different hubs depending on the destination. He develops an enumeration method for the multiple allocations and a branch-bound method for single allocation cases. P. Jaillet *et al.* (1996) proposed a different approach for the design of airline networks. First, they didn't assume an HS structure, and didn't consist of locating a given fixed number of hubs. Second, they modeled to track the number of passengers on a given flight, and to involve the choice of different aircraft

2000 *Mathematics Subject Classification.* 62K15, 65C50, 68T20 .
Key words and phrases. hub network design, multicommodity flow, dual based method, air cargo systems .

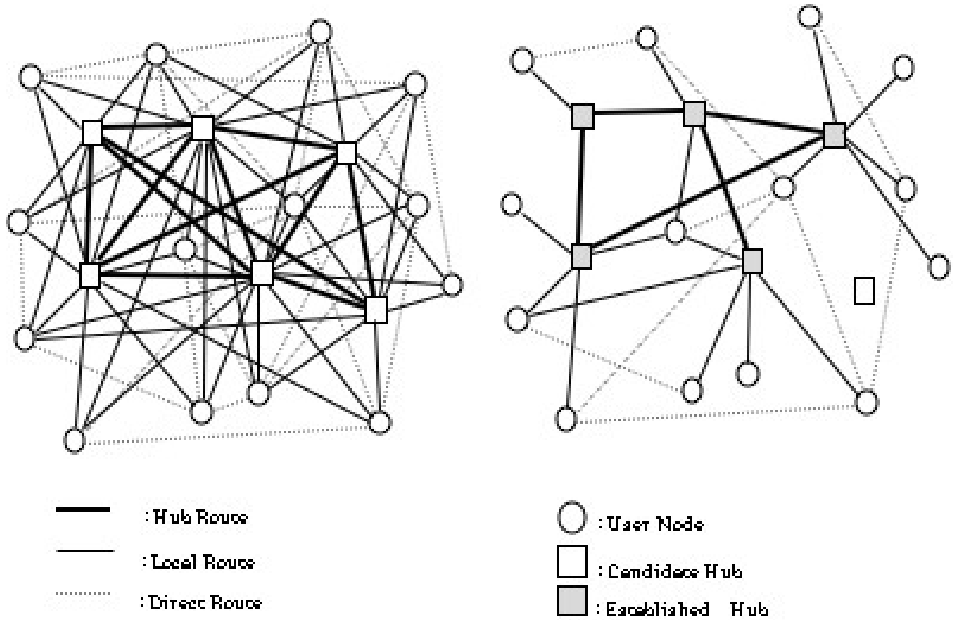

Figure 1: An example of hub-and-spoke network design problem

types of capacity and of the number of aircraft of each type to meet the demand. Third, their models allow many different paths between a given o-d pair. These three points considerably change the nature of the problem, and a direct quantitative comparison with the previous models would be meaningless. Horner and OfKelley (2001) also considered a similar problem with Jaillet *et al.* (1996). They did not consider an HS structure a priori as well, but modeled the level of link discount as a non-linear function of link flow volume. Sasaki *et al.* (1999) considered the 1-stop multiple p-hub median problem which is suitable for domestic airline network in a relatively small country. They formulated the model as a 0-1 Integer programming problem, which may further be transformed into a p-median problem. A branch-and-bound algorithm and a greedy-type heuristic were suggested to solve their problem. Mayer and Wagner (2002) focused on the uncapacitated multiple allocation hub location problem, which consists of locating hubs and determining the spokes between non-hub nodes and hubs. They developed an optimal algorithm by using a branch and bound procedure. Some location and network design problems are closely related with the HS network design problems. They include Campbell (1994, 1996), Melkote and Daskin (2001), OfKelly *et al.* (1996) and Smith *et al.* (1996).

Most of all papers on HS networks in airlines have focused on passenger travel. In the conventional passenger model, it is usually assumed that trips enable two stops, which is very useful in dealing with real-world HS networks. The two stop model allows, at most, two-hub stop services for each passenger travel, because more stops degrade the quality of service. However, in cargo systems, the importance is on time and safe delivery of their packages rather than the number of stops and the flight-route. Therefore, the damage/loss and the delay have been considered as the key quality parameters in the cargo services. Usually, a freight can be transferred from its origin to the destination via multiple airports within a delivery time. However, since a freight being shipped can be delayed at hub airports to consolidate and/or to wait an available flight, the frequent stops may result in a longer delivery time and increase the missing and/or the damaging opportunities of each package. That means the number of stops, i.e., hop-count, is the major factor affecting the quality parameters. Hop-constraints were already discussed in Balakrishnan and Altinkemer (1992) and Gouveia (1996) for representing constraints on a wide range of cost and service levels in network design problems. Therefore, it is desirable to add the additional restriction on the number of stops to increase the possibility of on-time and safe delivery. These points in cargo systems considerably change the nature of the conventional HS network design problem. In this paper, we focus on modeling the HS network design problem for air-cargo systems with hop-count constraints.

Implementing HS networks is generally attractive to airlines because of the cost saving derived from concentrating flow density on network links between hub nodes. When constructing an HS airline network in the conventional studies, two types of cost components have been included in the model: the fixed cost of establishing a hub airport, and the variable cost for transporting passenger on a flight-route. However, the movement of freight in HS air-cargo networks can be divided into two parts: Inter-hub links usually denote long distance transportation by large-scale cargo planes, and local links between hubs and local airports represent short distance transportation with small planes. When operating a cargo plane on a flight route, two types of costs should be considered: the fixed cost to set-up a flight route and the variable cost for transporting packages on the route. If a cargo aircraft is assigned to transport packages on a flight-route, the fixed cost for the correspondence to set-up and to maintain the cargo aircraft should be considered on that arc, and the packages can be transported by cargo aircraft on the route. On the other hand, if no cargo aircraft is needed on a certain route, no packages can be transported on the route. Therefore, for when designing an HS air-cargo network, the three major costs that should be considered in the optimization model are: the fixed cost of establishing a hub, the fixed cost to set-up a link, and the variable cost to transport freights on the link. The complex interrelationships between the cost elements make it extremely difficult to find the optimal HS network design where all three costs are optimally trade off. This may explain why it is hard to find a study in the literature which covers all three costs in a single optimization model. In this paper, with three major costs, we will develop an optimization model for a HS air-cargo network design problem with hop-count constraints. within a single framework, and suggest an efficient heuristic for our complex problem.

Our problem is now specifically described: (1) The site location of the local and the candidate hub airports which correspond to the user and the hub nodes respectively is given, and arcs representing flight routes between two airports are given. Also known is the demand for each origin-destination pair of user nodes. (2) We assume there is no limit on the capacity of an arc, or on the number of user nodes assigned to a hub, and .the total demand of a given origin-destination pair will be served via a single path only. (3) Each origin-destination demand can be served via hub nodes (i.e., via a hub path), or via a direct path between the origin and the destination nodes which does not pass through a hub node. However, the hub path should have a hop count limit. We assume that each user node may be connected to multiple hub nodes. (4) Three major cost elements are considered: the fixed costs of establishing the hubs, the fixed costs of including arcs in the network, and the variable costs associated with the arcs to satisfy demand. The problem to determine is: (1) the location and the number of hubs to be established, (2) which hub-hub and hub-user arcs should be included, (3) the routes used to satisfy O-D demands in such a way as to minimize the total network cost.

The rest of this paper is organized as follows: The next section describes an optimization model represented by a 0-1 Integer Programming (IP) formulation. Exploiting the model structure, we apply a dual ascent method to solve our model efficiently in section

3. Computational experiments are represented in section 4 to test our solution method for the sample network and some randomly generated problems by CPLEX program. Some concluding remarks and further researches on the solution methods are mentioned in the last section.

2 Model Formulation Consider an undirected network $G = (N, E)$ where N and E represent a set of nodes and edges respectively. N consists of the set of user nodes (local airports) I and the set of hub nodes (candidate hub airports) J. We define the set of directed arcs A by associating each undirected arc in E with two directed arcs having opposite directions. In order to discriminate arc types, the undirected and directed arcs are represented as $\{i, j\}$ and (i, j) respectively.

For our complex network design problem, we shall use a multicommodity flow formulation which has a suitable problem structure for developing an efficient solution method. Each origin-destination demand corresponds to an individual commodity k, and $o(k)$ and $d(k)$ denote its origin and destination nodes respectively. r_k denotes the amount of demand to be transported from $o(k)$ to $d(k)$. Let K be the set of those commodities.

To facilitate the problem formulation, consider the following notations:

z_j: the 0-1 variable concerning the establishment of a candidate hub airport j,

y_{ij}: the 0-1 variable concerning the use of arc $\{i, j\}$,

x_{ij}^k: the variable denoting the demand fraction of commodity k transported on arc (i, j),

g_j: the fixed cost incurred to establish a hub at candidate site $j \in J$,

f_{ij}: the fixed cost incurred to use an edge $\{i, j\}$ in E,

c_{ij}^k: the variable cost required for the demand of commodity k on an arc (i, j), which is set equal to c_{ji}^k.

With these notations, we now present the multicommodity flow model for our design problem.

$$(1) \quad (\mathbf{P}) \quad Min. \quad \sum_{j \in J} g_j z_j + \sum_{\{i,j\} \in E} f_{ij} y_{ij} + \sum_{k \in K} \sum_{(i,j) \in A} c_{ij}^k x_{ij}^k,$$

$$(2) \qquad s.t. \quad \sum_{j \in N} x_{ij}^k - \sum_{j \in N} x_{ji}^k = \begin{cases} i = o(k), \\ -1, i = d(k), \\ 0, otherwise, \end{cases} \quad i \in N, k \in K,$$

$$(3) \qquad x_{ij}^k \le y_{ij}, \{i, j\} \in E, \quad k \in K,$$

$$(4) \qquad x_{ji}^k \le y_{ij}, \{i, j\} \in E, \quad k \in K,$$

$$(5) \qquad \sum_{i \in N} x_{ji}^k \le z_j, j \in \ J, \quad k \in K,$$

$$(6) \qquad \sum_{i \in N} \sum_{j \in N} x_{ij}^k \le h_k, \qquad k \in K,$$

$$(7) \qquad z_j \in \{0, 1\}, \, y_{ij} \in \{0, 1\}, \, x_{ij}^k, \, x_{ji}^k \ge 0, \, j \in J, \, \{i, j\} \in E, \, k \in K,$$

The objective function of (P) has three cost terms: the hub establishment costs, the fixed costs of using arcs, and the variable costs on the arcs to transport demand. The fixed cost for hub j contains the hub establishment cost and the fixed part of all operating costs at hub j including cargo handling cost. The fixed cost on an arc (i, j), is represented as the fixed part for set-up, maintain and operating costs on arc (i, j). The variable cost on an arc (i, j) includes all kinds of variable costs occurring on arc (i, j) and node (airport) i such as variable costs for transporting on arc (i, j) and for handling cargo at airport i. The flow conservation constraints (2) enforce the network connectivity for each commodity. The flow restrictions of (3) and (4) force that flow on the arc be allowed in both directions only if the arc is used. Constraints (5) denote that the flow for each commodity can be shipped only on the arcs incident to the established hub node. The arcs incident to hub node i can be used only if the hub airport i is opened. Constraints (6) indicate the hop-count constraints. Each demand for commodity k should be reached from $o(k)$ to $d(k)$ via the path constituted by the number of arcs within a hop-count limit h_k. Constraints (6) are flexible in that it can be used to represent several quality constraints such as delivery time, damage and loss rates in cargo networks. For example, if every package have a delivery time limit (T_k) from the origin to the destination, it can be represented as the following constraint by using (6):

$$\sum_{i \in N} \sum_{j \in N} b_{ij}^k x_{ij}^k \;\le\; T_k, \qquad k \in K,$$

where, b_{ij}^k denote the travel time on the arc (i, j) and the ground handling time at the airport i including loading/unloading and waiting time for the next available flight. If every package has a limit for the damage and loss rate from the origin to the destination, it also can be represented as the following constraint:

$$\sum_{i \in N} \sum_{j \in N} a_{ij}^k x_{ij}^k \;\le\; l_k, \qquad k \in K,$$

where, $a_{ij}^k = ln(1 - m_i^k)$ and m_i^k denotes the damage and loss rate at the airport i for package k. $l_k = ln(1 - L_k)$ and L_k denotes the maximum allowance for the damage and loss rate for the package k. Therefore, the hop-count constraints can be expanded to represent various of quality constraints. If the hop-count constraint can be handled effectively in the solution procedure, other constraints based on the hop-count may be resolved with easy.

To consider a non-hub path for a user node pair in (P), the variable cost on the arc connecting the user nodes, which does not directly connect $o(k)$ to $d(k)$, should be set infinite: $c_{ij}^k = \infty$, $i \neq o(k) \in I$ or $j \neq d(k) \in I$. With this cost definition, we guarantee that no user node (local airport) can be used as a transit node for other commodities.

Although the model (P) is similar to the existing network design model, it is a 0-1 Integer Programming model and still has a comprehensive form of a network hub location problem which contains three kinds of decision variables: x's, y's and z's. If there are no restrictions for the number of intermediate nodes on the flight path for each commodity, the model is the same as that of the hub location and network design problems (Yoon *et al.*, 2000, Yoon and Current, 2003) which have NP-hard computational complexity. This study may then be viewed as the extended version of the one by Yoon and Current (2003). Exploiting and modifying their algorithm, we will develop a dual-based heuristic for solving our complex problem.

3 Solution Methods The problem (P) is a 0-1 IP problem, which contains a hub location problem as well as a network design problem with hop-count constraints. When we relax constraints (5) and (6), the remaining problem is an uncapacitated network design problem having NP-hard computational complexity (Balakrishnan *et al.*, 1989, Magnanti and Wong, 1984). Owing to the problem complexity, it is more effective to get a good feasible solution by a heuristic algorithm than to find an optimal solution. In this paper, we develop an heuristic algorithm to solve our comprehensive problem with ease.

Our heuristic solution method has two basic stages. The first, a dual ascent heuristic, solves a dual formulation of the LP relaxation of (**P**) to obtain a lower bound on the optimal solution to (**P**). Note that a dual ascent heuristic generates a good feasible solution within a short computation time, which can't guarantee an optimal solution for a dual formulation of the LP relaxation. The second stage uses information obtained from the dual solution to identify a feasible solution to the primal problem (**P**). Consider the dual of the LP relaxation of (**P**), where all the 0-1 variables are relaxed into non-negative variables.

$$(8) \quad (\mathbf{D}) \quad Max. \quad \sum_{k\in K} v^k_{o(k)} - \sum_{k\in K} h_k t_k,$$

$$(9) \quad s.t. \quad s^k_{ij} = \tilde{c}^k_{ij} - v^k_i + v^k_j + t_k \geq 0, \quad (i,j)\in A,\, k\in K,$$

$$(10) \quad s_{ij} = f_{ij} - \sum_{k\in K}(w^k_{ij} + w^k_{ji}) \geq 0, \quad \{i,j\}\in E,$$

$$(11) \quad s_j = g_j - \sum_{k\in K} u^k_j \geq 0, \quad j\in J,$$

$$(12) \quad w^k_{ij},\, w^k_{ji},\, v^k_i,\, u^k_j,\, t_k \geq 0, \quad \{i,j\}\in E,\, i\in I,\, j\in J,\, k\in K.$$

where,

$$\tilde{c}^k_{ij} = \begin{cases} c^k_{ij} + w^k_{ij} + u^k_i, & i\in J,\, (i,j)\in A,\, k\in K, \\ c^k_{ij} + w^k_{ij}, & otherwise. \end{cases}$$

The dual variables v^k_i's correspond to constraint set (2), the w^k_{ij}'s to constraint sets (3) and (4), u^k_j's to constraint set (5) and t_k's to constraint set (6). For brevity, s^k_{ij}, s_{ij}, s_j will be referred to as a commodity slack, arc slack and hub node slack, respectively. The arc having $s_{ij} = 0$, and the hub node having $s_j = 0$ will be referred to as the zero-slack arc and the zero-slack node respectively. For each $k \in K$, one of the constraints (2) is redundant. That means one of dual variables for each $k \in K$ corresponding to the constraints (2) is meaningless. Thus, we arbitrarily set the dual variables $v^k_{d(k)} = 0$ (Balakrishnan *et al.*, 1989).

Let **u**, **v**, **w** and **t** be vectors of u^k_j, v^k_i, w^k_{ij} and t_k values respectively. Given **u**, **w** and **t**, (**D**) can be decomposed into a sub-problem for each commodity $k \in K$ which denotes the dual of the shortest path from $o(k)$ to $d(k)$ with arc lengths $\tilde{c}^k_{ij}$'s (Magnanti and Wong, 1984, Balakrishnan *et al.*, 1989). Since the objective function of each sub-problem, $v^k_{o(k)}$, denotes the length of the shortest path, the main objective is how to increase $v^k_{o(k)}$ for each $k \in K$ by manipulating **u**, **w** and **t** effectively.

3.1 Dual Ascent Procedure Given **t**, (**D**) becomes none other than the LP relaxed dual of the hub location and network design problem studied by Yoon and Current (2003). They suggested an efficient dual ascent method to increase $v^k_{o(k)}$ by adjusting **u** and **w** for solving their problem, which is based on the labeling dual ascent method developed by Balakrishnan *et al.*(1989). In this paper, at each iteration with updated **t**, we simply apply the dual ascent method developed by Yoon and Current (2003), which improves the dual objective value without violating the dual feasibility by manipulating **u**, **v**, and **w**. The dual objective value may be further improved by increasing **t** iteratively in a way that does not violate the dual feasibility.

Given **t**, in order to increase the shortest path length $v^k_{o(k)}$ for a commodity $k \in K$, we introduce a brief sketch of Yoon and Current's dual ascent method (2003). With given **t**, we first identify directed arcs (i,j) whose , v^k_i, w^k_{ij} and/or u^k_j values should be increased. Let $A(k)$ be a set of those arcs. The nodes in **N** are divided into the two node sets **N**(k)

and $\mathbf{N}$'(k) by arcs in $A(k)$. Let $\mathbf{N}(k)$ denote the set of nodes connecting the origin of k, and $\mathbf{N}$'(k) be the set of remaining nodes. Then $A(k)$ can be divided into three types of arc sets:

$$\begin{aligned} A_1(k) &= \{(i,j) \in A : s_{ij}^k > 0\,, i \in N(k),\, j \in N'(k)\}, \\ A_2(k) &= \{(i,j) \in A : s_{ij}^k = 0,\, s_{ij} > 0,\, i \in N(k),\, j \in N'(k)\,\}, \\ A_3(k) &= \{(i,j) \in A : s_{ij}^k = 0,\, s_{ij} = 0,\, i \in N(k),\, j \in N'(k), s_i > 0,\, i \in J\,\}. \end{aligned}$$

Let $\mathbf{HUB}(k)$ be the set of hub nodes connected by arcs in $A_3(k)$: $\mathbf{HUB}(k) = \{\, i \in J : (i,j) \in A_3(k)\}$.Once $\mathbf{u}$ and $\mathbf{w}$ are adjusted, we can identify the arc type in $\mathbf{A}(k)$. If there exists any arc (i,j) in $A(k)$ having $s_{ij}^k = 0$, $s_{ij} = 0$, $s_i = 0$, $i \in J$, or $s_{ij}^k = 0$, $s_{ij} = 0$, $i \notin J$, the node j is deleted from $\mathbf{N}$'(k) and added to $\mathbf{N}(k)$, $A_1(k)$, $A_2(k)$ and $A_3(k)$ are updated. Thereby, the dual objective value can be increased by adjusting, v_i^k, w_{ij}^k and/or u_j^k values on the arcs in $A(k)$.

Let the amount of maximum increase of $\tilde{c}_{ij}^k$ on the arcs in $A_1(k)$, $A_2(k)$ and $A_3(k)$, be Δ_1, Δ_2, and Δ_3 respectively: $\Delta_1 = \min\{\, s_{ij}^k > 0 : (i,j) \in A_1(k)\}$, $\Delta_2 = \min\{\, s_{ij} > 0 : (i,j) \in A_2(k)\}$, $\Delta_3 = \min\{\, s_i > 0 : (i,j) \in A_3(k)\}$. The increases of $\tilde{c}_{ij}^k$ are bounded by constraints (9), (10) and (11) to maintain the dual feasibility; therefore, the maximum increase of $\tilde{c}_{ij}^k$ is determined by $\Delta = \min\{\, \Delta_1,\, \Delta_2,\, \Delta_3\}$. If $\tilde{c}_{ij}^k$ are increased by Δ on $A(k)$, then the dual objective value $v_{o(k)}^k$ will be increased by Δ, and the corresponding slacks would be adjusted on the arcs in $A(k)$. Special note should be focused on the nodes in $\mathbf{HUB}(k)$, $k \in K$. Consider a hub node $i \in \mathbf{HUB}(k)$ where at least one of the incident arcs of node i is in the arc set $A_3(k)$. From (9) and (10), we can see the increase of u_i^k without violating constraints (11) results in an increase of $\tilde{c}_{ij}^k$ for every arc incident to node i without adjusting w_{ij}^k. It provides a chance of an increasing of $v_{o(k)}^k$ for the commodity k. This series of ascent operations is continued until the resulting dual objective value is not increased.

Given $\mathbf{t}$, the procedure is described as follows :

uw-Procedure

Step 0 Set $\mathbf{CAND} = \mathbf{K}$

Step 1 Select $k \in \mathbf{CAND}$, and set $\mathbf{N}(k) = \{o(k)\}$ and $\mathbf{N}$'$(k) = N \backslash N(k)$.

Step 2 Define $A_1(k)$, $A_2(k)$ and $A_3(k)$, and the set of hub nodes in $\mathbf{N}(k)$,
HUB(k): HUB (k) = N(k)∩J.
Calculate the increase Δ_1, Δ_2, and Δ_3, and set $\Delta = \min\{\, \Delta_1,\, \Delta_2,\, \Delta_3\}$.

Step 3 Update the dual variables and slacks:
a. Adjust $v_i^k : v_i^k \leftarrow v_i^k + \Delta$, $i \in \mathbf{N}(k)$.
b. If $\Delta = \Delta_1$, $s_{ij}^k \leftarrow s_{ij}^k - \Delta$, $(i,j) \in A_1$, $and\, i \notin \mathbf{HUB}(k)$.
If $\Delta = \Delta_2$, $w_{ij}^k \leftarrow w_{ij}^k + \Delta$, $s_{ij} \leftarrow s_{ij} - \Delta$, $(i,j) \in A_2$.
If $\Delta = \Delta_3$, $u_i^k \leftarrow u_i^k + \Delta$, $i \in \mathbf{HUB}(k)$, $s_i \leftarrow s_i - \Delta$, $i \in \mathbf{HUB}(k)$.

Step 4 If $s_{ij}^k = s_{ij} = 0$, $i \in \mathbf{N}(k) \cap I$, $j \in \mathbf{N}$'(k), or $s_{ij}^k = s_{ij} = s_i = 0$, $i \in \mathbf{N}(k) \cap \mathbf{J}$, $j \in \mathbf{N}$'(k), then label the node j and add it to $\mathbf{N}(k)$. Let $\mathbf{N}(k) \leftarrow \mathbf{N}(k) \cap \{j\}$ and update $\mathbf{N}$'(k).

Step 5 If $d(k)$ is labeled, delete k from $\mathbf{CAND}$. If $\mathbf{CAND} = \emptyset$, go to **Step 6**.
Otherwise, return to **Step 1**.

Step 6 Stop.

The dual objective value can be improved further by adjusting $\mathbf{t}$ values. When updating

t, required then is keen awareness of two counting effects on the dual objective value: one is from the second term of (7) that the dual objective value is strictly decreased by increasing **t**. The other is that the increase of **t** is followed by the increase of $\tilde{c}_{ij}^k$s in (9), and hence of some $v_{o(k)}^k$.

Once **uw-procedure** is terminated, we have a set of arcs **CANE(k)** for each commodity $k \in K$ in which the increase of dual objective value can be achieved by adjusting **t**: $\mathbf{CANE}(k) = \{(i,j) \in A : s_{ij}^k = s_{ij} = 0, i \notin J \text{ or } s_{ij}^k = s_{ij} = s_i = 0, i \in \mathbf{J}\}$. For a commodity $k \in K$, we select the path which has a minimum hop-count from $o(k)$ to $d(k)$ in $\mathbf{CANE}(k)$. Let $\mathbf{P}(k)$ and $h(k)$ be the set of arcs and the number of arcs in the path. If we increase t_k by Δ where $h(k)$ is less than or equal to h_k, **(D)** reveals that the sum of maximum increment of s_{ij}^k on the path can not exceed the decrease of the objective value $\Delta \cdot h_k$. Thus, the objective value can be improved by increasing t_k only if $h(k)$ is greater than the hop-count limit h_k. Let $N_p(k)$ denote the set of nodes included in $\mathbf{CANE}(k)$ and $N_p'(k)$ be the rest of nodes in $\mathbf{N}$. If $h(k) > h_k$, we can consider the increase of t_k which is bounded by the dual feasibility. Calculate $\Delta_1(k)$, $\Delta_2(k)$, and $\Delta_3(k)$ as follows:

$$\begin{aligned}
\Delta_1(k) &= \min\{\, s_{ij}^k \,:\, s_{ij}^k > 0,\ (i,j) \in A\,\},\\
\Delta_2(k) &= \min\{\, s_{ij} \,:\, s_{ij}^k = 0,\, s_{ij} > 0,\ (i,j) \in A\,\},\\
\Delta_3(k) &= \min\{\, s_i \,:\, s_{ij}^k = 0,\, s_{ij} = 0,\, s_i > 0,\ i \in J,\, (i,j) \in A\,\}.
\end{aligned}$$

To maintain a dual feasibility, the maximum increase of t_k is determined by $\Delta(k) = \min\{\,\Delta_1(k), \Delta_2(k), \Delta_3(k)\}$. Once t_k is updated by $\Delta(k)$, **uw-procedure** is applied to improve $v_{o(k)}^k$ by adjusting s_{ij}^k, s_{ij} and s_i. The **t-adjustment** phase is terminated when the hop-counts of the shortest paths for all commodities are lesser than or equal to their hop-count limits. The formal procedure of **t-adjustment** phase is as follows:

t-adjustment

Step 0 Set $\mathbf{CAND} = \mathbf{K}$.
Step 1 Select $k \in \mathbf{CAND}$, and define $\mathbf{CANE}(k)$.
Step 2 Find the path $\mathbf{P}(k)$ having the minimum hop-count and $h(k)$ in $\mathbf{CANE}(k)$.
Find $N_p(k)$ and N_p' from $\mathbf{P}(k)$.
Step 3 If $h(k) \leq h_k$, delete k from **CAND** and go to **Step 5**.
f $h(k) > h_k$, calculate $\Delta_1(k)$, $\Delta_2(k)$, $\Delta_3(k)$ and update t_k:
$t_k \leftarrow t_k + \Delta(k)$, $(\Delta(k) = min\{\Delta_1(k), \Delta_2(k), \Delta_3(k)\})$.
Step 4 Applying Dual Ascent Heuristic with updated t_k to increase the dual objective value.
Step 5 If $\mathbf{CAND} = \emptyset$, then stop. Otherwise, return to **Step 1**.

3.2 Primal Heuristic Procedure Close inspection of the dual ascent heuristic reveals that the dual solutions generated from this algorithm give us some important information. At every step in the algorithm, v_i^k for all $i \in N$ and $k \in K$ represents the shortest path length from the origin to the destination node using the arc lengths $\tilde{c}_{ij}^k$. When the dual ascent procedure terminates, for every commodity $k \in K$, there is at least one path from the origin to the destination within the hop-count limit. (If otherwise, the dual ascent heuristic could not terminate). Furthermore, the path contains only zero-slack arcs. This information provides a clue to construct a feasible network design.

To find a good primal solution, we first construct an initial feasible solution for (P) by using the dual solution, and improve it by eliminating the unnecessary or the redundant nodes and arcs. For that we consider the complementary slackness (CS) conditions for the LP relaxation problem of (P). Even (P) is Integer Programming model, its feasible solution obtained from an heuristic will be close to the optimal solution, once the violation of CS conditions for the LP relaxation problem of (P) is minimized. The objective of our primal heuristic is to obtain a set of primal feasible variables $\{x_{ij}^k\}$, $\{y_{ij}\}$, $\{z_j\}$ which, with the dual solution at hand, satisfies the CS conditions as much as possible.

Let consider the following CS conditions which are based on the dual problem (D) for the LP relaxation problem of (P):

$$s_{ij}^k \cdot x_{ij}^k = 0, \quad (i,j) \in A, \tag{13}$$

$$s_{ij} \cdot y_{ij} = 0, \quad \{i,j\} \in E, \tag{14}$$

$$s_j \cdot z_j = 0, \quad j \in J. \tag{15}$$

In order to satisfy these CS conditions, we can find that primal feasible variables have to take non-zero values on the zero-slack arcs and the zero-slack nodes only. We will select the set of zero-slack arcs and the set of zero-slack nodes as the candidate arcs and hub nodes being included in a feasible solution. Let E^* and J^* denote those arc and node sets respectively:

$$E^* = \{\{i,j\} : s_{ij} = 0, \{i,j\} \in E\}, \; J^* = \{j : s_j = 0, j \in J\}$$

E^* may have some unnecessary arcs. Since an arc incident to a hub j can be used only if the hub j is established, the arcs incident to the nodes in $J' = J \backslash J^*$ are eliminated from E^*. For a commodity $k \in K$, A_k^* denotes the set of arcs having zero commodity slacks on A^* obtained from E^*. We can find the shortest path with hop-count limit h_k for k using c_{ik}^k as arc lengths on A_k^*. To find the shortest path, we apply the Bellman-Ford algorithm (Nemhauser and Wolsey, 1988) which generates the shortest path from origin to all nodes within a certain iteration. Let A_k^+ be the set of arcs on this shortest path. A_k^+ can be used to define the set of established hubs J^+ from J^* as follows: $J^+ = \{i : (i,j) \in A_k^+, i \in J^*, k \in K\}$. From the network $G^+ = (I \cup J^+, E^+)$, where $E^+ = \{\{i,j\} \in E^* : (i,j) \in A_k^+, or (j,i) \in A_k^+, k \in K\}$, the primal feasible solution is then constructed by the following:

$$x_{ij}^k = \begin{cases} 1, & (i,j) \in A_k^+, \\ 0, & Otherwise, \end{cases} \quad k \in K, \quad y_{ij} = \begin{cases} 1, & (i,j) \in E^+, \\ 0, & Otherwise, \end{cases} \quad z_j = \begin{cases} 1, & j \in J^+, \\ 0, & Otherwise. \end{cases}$$

Although this heuristic generates a feasible solution to the primal problem (**P**), G^+ may contain more arcs and hub nodes than necessary. By eliminating unnecessary hubs and arcs, we can reduce the total cost. A drop type heuristic is applied to remove unnecessary hubs and arcs from G^+ one by one until no cost reduction can be made from such an arc or a node-drop (Balakrishnan *et al.*, 1989, Yoon *et al.*, 1998). We apply the node-drop heuristic first, and then the arc-drop heuristic to remove unnecessary hub nodes and arcs from G^+. The node(arc)-drop heuristic starts with an initial solution, and at each iteration drops a hub node(arc) from the current feasible solution to reduce the total cost. The procedure continues until no node(arc) can be dropped from the solution without increasing the total cost.

4 Computational Experiments The solution procedure was coded in C, and test runs were performed on a PC(Pentium IV/1GHz) to evaluate the quality of heuristic solutions.

We first tried to find optimal solutions of (**P**) for the sample network with 6 candidate hub airports and 14 local airports (Figure 2) by using CPLEX program(ILOG, 2002), and solved 36 randomly generated problems by using our heuristic.

To generate various cost data, we followed the procedure as done in Yoon *et al.* (2000). We first set a base cost c_{ij} on each arc $\{i, j\}$, which was the Euclidean distance between two nodes i and j. The fixed cost of arc $\{i, j\}$ was obtained by multiplying the base cost by the scaling factor f which was the same for all arcs. The variable cost of each commodity, k, on arc (i, j) was obtained by multiplying the base cost c_{ij} by the demand U_k. Each demand U_k, $k \in K$, was randomly selected from an interval (10.0, 50.0). The hub establishment costs, g_j, were chosen randomly from interval (a, b).

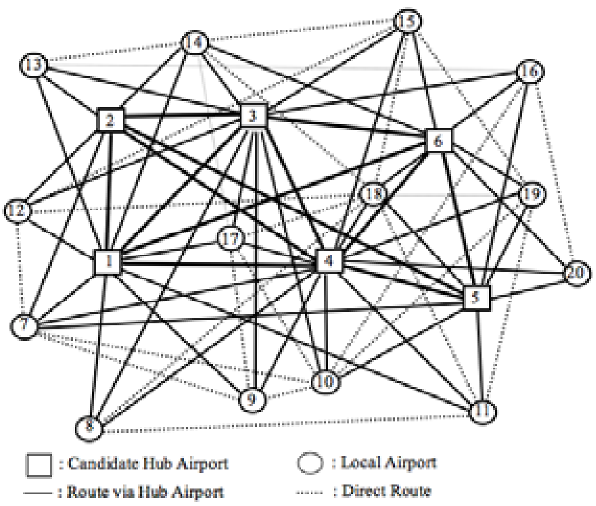

Figure 2: An Example Network

To investigate the variations of hub-and-spoke networks according to the hop-count limit, we consider the two types of hop-count constraints for the sample network. One is a 3-hops limit which is commonly applied for all commodities. The other is the mixture of 2 and 3 hops. that is, the half of commodities has a 2-hops limit and the remaining commodities have a 3-hops limit. We attempt to solve all of the sample problems optimally by using CPLEX program. Table 1 lists the summary of the test problems and the computational results.

In networks where the fixed costs on the arcs are relatively low, each local airport has multiple hub airports to transport their demand with minimal cost. As the fixed cost on the arc becomes relatively high, the number of flight-routes (arcs) is decreased, but the number of routing-paths via multiple hubs is increased to save on transportation costs. The number of direct or non-hub connections increase as the hub fixed costs increase. Another notable one is that the number of direct routes being included in the network increase sensibly with the hop-count constraints. In Table 1, we can find the hop-count constraint is a very critical factor in the hub-and-spoke system. Once the half of commodities is restricted to have 2-hops limit, the number of hub airports being established and the number of flight-routes required to transport commodities become more than that of having 3-hops limit for all commodities, and the objective value is increased by 300 times.

To test our method for more general cases, we generate the test problems randomly,

Table 1: Input parameters and computational result

g_j	f	h_k	Objectiv Value	No.of hubs	No. of arcs Hub	Spoke	Direct	Total	No. of routing paths 1Hub	2Hub	Hub0	Times (sec.)
[1,000 - 3,000]	10	3	19,164	2	1	18	3	22	66	22	3	1.95
	100		24,335	2	1	15	2	18	50	39	2	3.25
	500		40,513	3	3	14	0	17	60	31	0	6.33
	10	2/3	6,021,313	4	4	32	9	45	73	9	9	1.70
	100		6,031,200	4	4	31	6	41	75	10	6	4.63
	500		6,070,233	4	1	28	6	35	76	9	6	4.82
[5,000 - 7000]	10	3	27,634	2	1	18	3	22	66	22	3	2.66
	100		32,335	2	1	15	2	18	50	39	2	3.23
	500		48,604	2	1	14	0	15	46	45	0	4.56
	10	2/3	6,037,313	4	4	32	9	45	73	9	9	1.68
	100		6,042,202	4	2	31	6	39	76	9	6	3.52
	500		6,086,233	4	1	28	6	35	76	9	6	8.99
[10,000 - 20,000]	10	3	44,634	2	1	18	3	22	66	22	3	18.02
	100		49,335	2	1	15	2	18	50	39	2	72.59
	500		65,604	2	1	14	0	15	46	45	0	7.12
	10	2/3	6,068,313	4	4	32	9	45	73	9	9	1.80
	100		6,078,195	4	4	31	6	41	75	10	6	4.24
	500		6,117,233	4	1	28	6	35	76	9	6	3.39

but systematically. We first randomly located the pre-specified number of candidate hub and user nodes on a (100 x 100) grid in a plane. We located a spanning tree covering all selected nodes. On top of this spanning tree, additional edges were randomly placed until the pre-specified number of edges was obtained. To guarantee a feasible solution, we made certain that each user node was connected to at least one hub node with an arc. The cost data in the general network are defined by the same way on the sample network. We consider a total of 36 randomly generated problems with a hop-count limit of 3, and tried to solve them by using CPLEX program and our dual-based heuristic.

The test problems were grouped into 4 different sets according to the network size. Each subset was divided into 9 problems by the cost parameters, i.e., the range of hub fixed costs and the scaling factor **f** on arcs. In Table 2, the optimal value(ZP*) is obtained by using CPLEX program for the model (**P**). The primal objective value(ZP), i.e. the best upper bound, and the dual objective value (ZD), the lower bound, are obtained by our dual-based heuristic. % gap is the ratio of difference between the best upper bound of the objective value and the lower bound for problem (**P**). The details of the input parameters and the computational results for test problems are given in Table 2.

For the first two subsets, even though the best upper bounds have small % gaps, they are almost the same as that of the objective values. This means that our heuristic either finds the optimal solutions or the tighter upper bounds for the first two subsets. However, considering the computation times, CPLEX takes more than 10 times for finding the optimal solution compared to our heuristic method. Although the comparison is performed on a small sized network, the computational results indicate that our heuristic generates good feasible solutions in reasonable time. For the last two subsets of test problems, we can not find the optimal solution because CPLEX can not treat such a large-scale problem on a PC. That is, the third set, 10 candidate hubs, 40 local nodes, 350 candidate flight routes and 780 commodities, has 546,360 variables, 593,580 constraints and 2,914,080 non-zero elements. Thus it is very difficult to find an optimal solution on such a large scale IP problem.

The percentage gaps between the best upper and lower bounds (duality gaps) for the smaller two problem sets are generally under 3% for 10 of 18 problems, and only two have an average above 6%. Gap averages for the larger problem sets range from 2.2% to 8.5%. Although appreciably larger than the smaller problems, they still compare favorably to results from other less complex problems reported in the literature (Jaillet et al., 1996, Sasaki et al., 1999, Yoon et al., 1998). The reader's attention is again called upon to the fact that our problem is so complex to include as subproblems in an integrated framework

Table 2: Randomly generated problems and computational results

Pro. Set	$\|I\| \times \|J\| \times \|E\| \times \|K\|$	[a,b]	f	Objective Value ZP*	ZP	ZD	% Gap	No. of Hub CP.	He.	No. of r.p. CP.	He.	Time(sec.) CP.	He.(dual)
I-1		[1,000	10	47,425	47,425	46,264	2.52	2	2	21	22	6.3	0.5
I-2		5,000]	50	50,459	50,492	49,443	2.12	2	2	21	21	6.3	0.4
I-3			100	54,252	54,317	53,620	1.30	2	2	21	21	6.6	0.2
I-4		[5,000	10	47,992	47,992	54,670	1.54	1	1	20	20	7.5	0.2
I-5	6x20x	10,000]	50	51,408	51,408	50,465	1.87	1	1	20	20	7.3	0.2
I-6	105x190		100	55,678	55,678	54,670	1.8	1	1	20	20	7.3	0.2
I-7		[10,000	10	53,094	53,094	52,592	0.95	1	1	20	20	7.4	0.2
I-8		20,000]	50	56,510	56,510	55,846	1.2	1	1	20	20	7.2	0.2
I-9			100	60,780	60,780	60,254	0.87	1	1	20	20	7.3	0.2
II-1		[1,000	10	93,835	93,945	89,334.5	5.1	4	4	57	57	292.1	33.0
II-2		5,000]	50	99,994	100,829	94,386.5	6.8	3	2	37	37	120.0	6.1
II-3			100	106,385	108,136	100,389	7.7	2	2	31	31	268.5	4.9
II-4		[5,000	10	111,576	111,576	100,515.5	5.7	2	2	37	37	69.1	5.7
II-5	10x30x	10,000]	50	117,093	117,110	110,869	5.6	2	2	31	31	117.3	17.7
II-6	300x435		100	123,455	123,490	117,392	2.2	2	2	31	31	111.9	13.7
II-7		[10,000	10	125,784	125,784	120,049	4.8	2	2	31	33	108.0	3.1
II-8		20,000]	50	130,859	130,876	126,214	3.7	2	2	31	31	99.6	4.3
II-9			100	136,921	136,956	133,048	2.9	2	2	31	31	80.4	3.4
III-1		[1,000	10	-	164,963	160,476.5	2.8	-	4	-	72	-	122.1
III-2		5,000]	50	-	178,647	166,703.5	7.1	-	3	-	59	-	156.1
III-3			100	-	189,452	174,647	8.5	-	3	-	53	-	149.7
III-4		[5,000	10	-	181,108	173,869	4.2	-	3	-	64	-	96.8
III-5	10x40x	10,000]	50	-	192,774	180,515.5	6.8	-	2	-	44	-	35.6
III-6	350x780		100	-	202,231	188,637	7.2	-	2	-	43	–	28.8
III-7		[10,000	10	-	199,585	185,301	7.7	-	2	-	47	-	16.6
III-8		20,000]	50	-	206,810	192,638	7.3	-	2	-	44	-	62.8
III-9			100	-	216,267	201,480	7.3	-	2	-	43	-	43.6
IV-1		[1,000	10	-	260,807	255,082.5	2.2	-	4	-	79	-	521.9
IV-2		5,000]	50	-	271,301	262,813.5	3.2	-	4	-	72	-	868.9
IV-3			100	-	284,219	272,169	4.4	-	3	-	57	-	105.9
IV-4		[5,000	10	-	254,291	245,194	3.7	-	3	-	65	-	82.6
IV-5	10x50x	10,000]	50	-	263,387	252,369	4.4	-	3	-	63	-	137.1
IV-6	600x1225		100	-	273,956	261,733	4.6	-	3	-	60	-	177.5
IV-7		[10,000	10	-	290,693	279,481	4.0	-	2	-	56	-	49.8
IV-8		20,000]	50	-	304,377	287,507	5.8	-	3	-	62	-	336.8
IV-9			100	-	315,643	297,963	5.9	-	3	-	59	-	346.5

Zip*: the optimal objective value obtain by CPLEX program.
ZP: the best primal objective value obtained by our heuristic. ZD: the dual objective value obtained by our heuristic.
%Gap = $(ZP - ZD)/ZD \times 100(\%)$
CP. : CPLEX, He. : Heuristic, r.p.: routing path

the two NP-hard problems: one on hub location and the other on design of an uncapacitated network with hop-count constraints. Despite the excessive computational burden expected for such problems, we can obtain a goon feasible solution for a large-sized network within a few minutes. Furthermore, noting that the best upper bound is very close to the optimal objective value for the smaller two sub sets, the primal objective values may be close to the optimal ones even though they have some duality gaps. The solution times are increased as the fixed costs of hub establishment and the fixed costs on arcs are increased. However, using our heuristic, most of problems are solved within a few seconds on a PC, and even for a large sized network, we can find a good feasible solution within a few minutes by using our heuristic.

5 Conclusion Hub-and-spoke topologies are frequently used in the design of transportation and communication networks because of their potential to reduce cost through economies of scale. Despite the many researches on hub-and-spoke network design problem in airlines, they have been focused on designing the network for passengers. Furthermore, only two types of costs, the fixed cost of establishing a hub and the variable cost on an arc for passenger traffic, are included in their models. In this paper, we addressed a hub-and-spoke network design problem for air-cargo systems. We considered the fixed costs of establishing the hubs and transportation arcs, as well as the variable costs of traversing these arcs. A freight being shipped can be delayed at hub airports to consolidate and/or to wait for an available flight. However, the excess delay makes the grade of service worse. In order to guarantee a certain level of delivery time, we consider the number of hop-counts which represents the number of hub airports being passed through. The problem is modeled as a variation of the multi-commodity network flow problem. Exploiting the model structure, we developed a dual-based heuristic, by which we can obtain a good feasible solution efficiently.

Computational experiments for test problems were conducted to show the satisfactory performance of the proposed heuristic.

Our model was tested on the sample network and the randomly generated networks of varying cost structures. Even the hop-count constraints make the problem complex, the results indicate that our model generates a good feasible solution favorably in a short computation time, and we can use it for the sensitivity analysis for various cost structure. With the computational experiments, our model can be expanded to more real-world air-cargo problems. Even, we concentrate on the full cargo network design problem in this paper, our model gives a fundament for modeling more practical problem including passenger-cargo combo flights.

References

[1] Aykin, T., 1995. Networking policies for hub-and-spoke systems with application to the air transportation system. Transportation Science 29, 201-221. Balakrishnan, A., Magnanti, T.L. and Wong, R.T., 1989. A dual-ascent procedure for large-scale uncapacitated network design. Operations Research 37, 716-740.

[2] Balakrishnan, A., Altinkemer, K., 1992. Using a hop-constrained model to generate alternative communication network design. ORSA Journal on Computing 4, 192-205.

[3] Bryan, D. and OfKelly, M.E., 1999. Hub and spoke networks in air transportation: an analytical review. Journal of Regional Science 39, 275-295.

[4] Campbell, J.F., 1994. A survey of hub location. Studies in Location Analysis 6, 31-49.

[5] Campbell, J.F., 1996. Hub location and the p-hub median problem. Operations Research 44, 923-935.

[6] Gouveia, L., 1996. Multicommodity flow models for spanning trees with hop constraints. European Journal of Operational Research 95, 178-190.

[7] Horner, M.W. and OfKelly, M.E., 2001. Embedding economies of scale concepts for hub network design. Journal of Transport Geography 9, 255-265.

[8] ILOG, 2000. CPLEX 8.0 Userfs manual, ILOG Inc.

[9] Jaillet, P. Song, G. and Yu, G., 1996. Airline network design and hub location problems. Location Science 4, 195-212.

[10] Magnanti, T.L. and Wong, R.T., 1984. Network design and transportation planning: models and algorithms. Transportation Science 18, 1-55.

[11] Mayer, G. and Wagner, B., 2002. HubLocator: an exact solution method for the multiple allocation hub location problem. Computers & Operations Research 29, 715-739.

[12] Melkote, S. and Daskin, M.S., 2001. An integrated model of facility location and transportation network design. Transportation Research Part A 35, 515-538.

[13] Nemhauser, G.L. and Wolsey, L.A., 1988. Integer and Combinatorial Optimization, John Willey & Sons, New York.

[14] O'Kelly, M.E., 1986. The location of interacting hub facilities. Transportation Science 20, 92-106.

[15] 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, 125-138.

[16] Sasaki, M. Suzuki, A. and Drezner, Z., 1999. On the selection of hub airports for an airline hub-and-spoke system. Computers & Operations Research 26, 1411-1422.

[17] Smith, K., Krishnamoorthy, M. and Palaniswami, M., 1996. On the location of interacting hub facilities: neural versus traditional approaches. Location Science 4, 125-138.

[18] Yoon, M.G., Beak, Y.H. and Tcha, D.W., 1998. Optimal design of a distributed fiber transport network with hubbing topology. European Journal of Operational Research 104, 510-520.

[19] Yoon, M.G., Beak, Y.H. and Tcha, D.W., 2000. Optimal design model for a distributed hierarchical network with fixed-charged facilities. International Journal of Management Science 6, 29-45.

[20] Yoon, M.G. and Current, J.R., 2003. The hub location and network design problem with fixed and variable arc costs: Formulation and dual based solution heuristic. Being revised for publication in Naval Logistics Research.

* Department of Business Administration Hankuk Aviation University200-1 Hwajun-dong, Koyang-shi, Kyunggi-do, Korea 412-791, Phone : +82-(0)2-300-0097
Email : mgyoon@hau.ac.kr

† Department of Business Administration Hankuk Aviation University 200-1 Hwajun-dong, Koyang-shi, Kyunggi-do, Korea 412-791, Phone : +82-(0)2-300-0097
Email : seomjingang@hotmail.com

‡ Department of Management information System Administration Nagoya University of Commerce & Business 4-4 Sagamine, Komenoki-cho, Nisshin-shi, Aichi-ken, Japan 470-0193, Phone : +81-(0)561-73-2111
E-mail : han-sh@nucba.ac.jp