Transportmetrica, Vol. 4, No. 1 (2008), 33-49

AIRLINE NETWORK DESIGN PROBLEM WITH DIFFERENT AIRPORT CAPACITY CONSTRAINTS

TA-HUI YANG[1]

Received 31 January 2007; received in revised form 30 May 2007; accepted 3 August 2007

Shortages of resources at airports have become a serious problem, however most past studies in airline network design have not taken into consideration capacity restrictions at airports. This study addressed the capacitated fixed charge hub location problems. Two models with different capacity constraints are proposed to determine the hub location and service network; capacity constraints in the first model take into account total volumes that pass through airports, including inbound, outbound, and transshipped traffic, while those in the second model only account for transshipment flows. Both models take into consideration most possible services, including non-stop and hub-connected, and the coexistence of both services. This provides maximal flexibility for the construction of airline service networks. A network based on the air freight market between Taiwan and China will be used to illustrate the proposed models, while analyses show that the use of different capacity constraints will result in different hub locations and network structures.

KEYWORDS: Airline network design, hub location, capacity limitation, mixed integer problem

1. INTRODUCTION

The problems of hub-and-spoke networks have drawn a great deal of attention in recent years. These networks have been widely applied in air transportation, telecommunications, express package deliveries, and other logistical systems. They have particularly dominated the air passenger, air freight, and parcel delivery businesses. The hub-and-spoke system pipes the traffic into a number of major facilities (called hubs) by way of "spokes". Transshipment is the shipment of flows to an intermediate airport, and then from there to another airport. In the hub-and-spoke system the intermediate stop has to be a hub airport. In practice, transshipments usually involve no more than two hub-stops. This operation simplifies network configuration by replacing large numbers of direct linkages between nodes with fewer transshipment connections. This simplification in network structure reduces network construction and operation costs.

Another benefit is that the carriers can take advantage of economies of scale through consolidation of the flows on the links between non-hub nodes and hubs, as well as the links between hubs. Channeling traffic into hubs makes it commercially viable to offer better quality services to a diverse range of origins and destinations. In hub-based networks, though travel distance and travel time are generally longer than non-stop flights, carriers usually increase schedule frequency to account for fewer service flight routes. In hub-and-spoke operations, passengers may be able to make a return trip in a day, but in the direct point-to-point services it would become an overnight stay which increases trip costs considerably. Detailed discussions of network economics for the hub-and-spoke structure can be found in Button (2002).

According to the fundamental concept of hub-and-spoke network models, there are two different formulations for analytical formulations: link-based and path-based formulations. In the link-based model, the decision variables are defined as the links between two distinct nodes. The link-based formulation is often used in typical network design problems, because it is usually composed of a more compact formulation.

[1] Department of Logistics Management, National Kaohsiung First University of Science and Technology, 2 Jhuoyue Rd., Kaohsiung 81164, Taiwan. Corresponding author (E-mail: yang@ccms.nkfust.edu.tw).

O'Kelly (1986a) first presented one-hub and two-hub location models in continuous planar space. Due to the nature of hub location applications, the choice of hub is often limited to a set of pre-defined sites. Therefore, a discrete version of this model could be more suitable for practical concerns. O'Kelly (1987) formulated discrete hub location problem as a quadratic integer program. The model was enhanced by replacing the objective function with three different criteria: minimizing transportation costs, minimizing total hub usage, and minimizing variability of hub usage (O'Kelly, 1987; O'Kelly, 1986b). Skorin-Kapov and Skorin-Kapov (1994) developed an algorithm based on tabu search to solve the quadratic integer hub location model proposed by O'Kelly (1987). In the above studies, the fixed cost of opening facilities is ignored and the total number of hubs to be constructed is exogenously given. O'Kelly (1992) introduced fixed facility costs into the hub location model and incorporated the number of hubs as decision variables. The major drawback of the link-based formulation is that it would result in integer problems with non-linear objective functions for multiple hub systems. Because of the absence of efficient solution algorithms for such problem, solving these real-world problems requires intense computational effort. Therefore, the optimal solution can only be reached for small systems.

In most real-world applications, it is generally necessary to explicitly maintain the path flows in the formulation. The path-based formulation uses the possible paths between every pair of origins and destinations as decision variables. Though the path-based models reduce the hub location problems from integer programs with non-linear objective functions to integer programs with linear objective functions, the benefits come at the cost of additional variables and constraints. Campbell (1994) proposed path-based formulations for four types of discrete hub location problem: p-hub media, uncapacitated hub location, p-hub center, and hub covering. Different methods and algorithms were then proposed to solve Campbell's models (Campbell, 1996; Klincewicz, 1996; O'Kelly et al., 1996; Skorin-Kapov et al., 1996; Sohn and Park, 1998). O'Kelly and Bryan (1998) adopted non-linear concave functions to represent the economies of scale on link costs. Because this problem is difficult to solve, a piece-wise linear model was proposed to approximate a solution. O'Kelly and Bryan (2002) investigated the behavior of four models in literature and concluded that reflecting the economic scale by concave cost functions favors the integer hub solution and discourages solution from fractional facilities.

The studies we have discussed so far limited the flows between every pair of nodes to shipping through hubs. However, in some applications, the flows may be permitted to be shipped directly to the destination without being transshipped through hubs. Aykin (1995a) proposed a model considering both non-stop service and hub-stop service. This model was further extended to a capacitated hub location model by considering capacity limitations at hubs (Aykin, 1994). Aykin (1995b) also applied the idea to hub-and-spoke network design in continuous planar space.

In hub-and-spoke systems, flows from different origins are concentrated into a number of hubs and shipped out to other hubs or to their destinations. Through the consolidation of link flows, traffic at the hubs is significantly increased. Therefore, hub capacity limitations become an important factor in hub network design. In the applications of air transportation, hub locations are usually determined from existing airports. This implies that the capability of handling the traffic is limited by the existing infrastructure on airports. Many major airports face serious shortages in airport resources, such as gates, terminal space, and takeoff/landing slots. To ease air traffic congestion, in 1969 the Federal Aviation Administration (FAA) began a quota system to limit the number of takeoffs and landings during peak periods at some busiest United States airports. Hence, capacity constraints cannot be ignored in the design of airline networks. However, few studies can be found addressing hub network design with capacity limitations. This study aims to explore the influence of capacity limitations on hub locations and network structure.

In addition, hubs might be equipped with different functions depending on application. Some are designed mainly as sorting or switching centers whose major traffic comes from transshipment flows, for example: Louisville, Kentucky for UPS and Memphis, Tennessee for FedEx. In some hubs, enplanements and deplanements contribute a significant portion of the total traffic in addition to transshipment flows. Therefore, hubs can be classified into two types of functions: 1) hubs designed for handling inbound, outbound, and transshipping traffic; and 2) hubs designed primarily for transshipping flows. Since the different types of functions require different facilities for traffic operation, resource limitations would be different for different types of hubs. Another objective of this paper is to investigate whether differences in the designed function of hubs might result in different hub locations and network structures.

One major contribution of this study is that the proposed models can determine the optimal network structure and flight routes. It does not only limit to hub-and-spoke system. The models might choose point-to-point operations, if they perform better. It is also possible to design a system with portions using point-to-point operations and others using hub-and-spoke operations. Furthermore, the models allow the coexistence of non-stop and hub-stop services and the coexistence of different hub-stop paths. This might not be necessary for incapacitated problems but is very important in the capacitated networks. In incapacitated problems, the demand for each origin-destination pair is always delivered via the lowest cost path. However, if the hub has a capacity limitation, the path with lowest cost might not be able to transport all the traffic to the destination because the path is constrained by the airport capacity. It may require more than one path (non-stop path plus hub-stop path or different hub-stop paths) to transport all demand for an origin-destination pair. This situation can be properly handled in the proposed models, since several paths between an origin-destination pair may simultaneously deliver the demand. In other words, the proposed models provide the most flexibility to airline network design, because the models are capable of taking all possible situations into account.

The remainder of the paper is organized as follows: Section 2 presents the problem description and formulations of airline network design under capacity limitations. Section 3 includes tests of the proposed models based on real data and sensitivity analysis. Section 4 is a summary and discussion.

2. PROBLEM DESRIPTION AND MODEL FORMULATION

Consider a system with a set of nodes to be served. In this study, the nodes are represented as the airports. Let N denote the set of nodes. d_{ij} represents the demand from node i to node j. The purpose of this study is to design a service network to transport the demands to the destinations at the lowest operation cost under capacity limitations at airports from the standpoint of an airline company.

In this study, a direct linkage between two airports is called a segment. A flight path (or route) represents the movement of flow from the origin to the destination. A non-stop flight path, which only contains one segment, represents the flow being directly shipped from origin to destination without any intermediate hub stops. A one-hub-stop flight path contains two segments and one intermediate stop on a hub. Similarly, a two-hub-stop flight path contains three segments and two intermediate stops on two different hubs.

Before we introduce the models, the following assumptions are made in order to reasonably reflect the reality and to clearly define the problem:

1. The airports to be served and their locations are known.
2. Every flight route is allowed at most two transshipments through different hubs. Three-hub-stop flights are rarely applied in practice. Therefore, this study does not consider the transshipment with more than two hub stops.
3. Three possible routing services are considered: direct flight (or non-stop flight), hub connected flight (one-hub-stop and two-hub-stop), and coexistence of both direct and hub connected flights.
4. The hubs are restricted to existing airports. We will not consider setting up a new hub in the planar space.
5. There is no barrier to acquiring rights of way between airports. In other words, the flight routes can connect to any airport if it is necessary for the airline company.
6. The freight is not perishable goods. Hence, the cost of freight deterioration is not considered in this study.
7. Hub-and-spoke operations enjoy economies of scale in transportation costs. The scale economies occur on the links between two hubs and the links between a non-hub node and a hub. They are approximated by different discount factors.
8. Airports are subject to capacity limitation. The capacity limitation at an airport for a particular airline company is related to many factors, including available gates, terminal space, landing and take-off slots … etc. Estimating the available capacity might not be an easy task. Since the purpose of this study is airline network planning, capacity estimation is not the focus here. Capacity is transferred to the equivalent volume of freight in tons, which is the maximum traffic allowed to be transported by an airline company at the airport. For example: Airline C is permitted to have 3 flights landing and take-off at Airport D every day and each plane can carry at most 100 tons cargo. If there are 260 working days per year, the annual maximum capacity for Airline C at Airport D is: (3 flights)$\times$(100 tons)$\times$(260 days) = 78,000 tons per year. This simplifies the procedure of estimating capacity limitations for an airline company at airports.

We use path-based formulation to model the network design problem. Therefore the decision variables are defined as the paths between origin-destination pairs. According to assumption 2, the flights are not allowed to stop more than two hubs. Hence, the variables and constraints have been dramatically reduced, though the path-based formulation is expected to have a large number of variables and constraints (as opposed to link-based formulation). The decision variables are defined as follows:

s_k : the auxiliary variable to represent the hub location. If $s_k = 1$, node *k* is set to be a hub. If $s_k = 0$, node *k* is not a hub.

x_{ij} : the fraction of flow transported by non-stop service from *i* to *j*. $0 \le x_{ij} \le 1$. By this definition, both nodes *i* and *j* are not hubs.

x_{iktj} : the fraction of flow from *i* to *j* and transshipped at hubs *k* and *t*. $0 \le x_{iktj} \le 1$.

Other notation and variables are defined as follows:

N : the set of all nodes. Each node represents an airport.

d_{ij} : the demand needing to be transported from *i* to *j*.

α : the discount factor for the inter-hub transportation costs. This approximates the economies of scale between two hubs. $0 \le \alpha \le 1$.

β : the discount factor for the transportation costs between a non-hub origin and a hub, or between a hub and a non-hub destination. It approximates the economies of scale between hubs and non-hub nodes. In general, $0 \le \alpha \le \beta \le 1$.

c_{ij} : the unit transportation cost for the non-stop service between *i* and *j*.

c_{iktj} : the unit transportation cost for hub-connected service between *i* to *j* and with transshipment at hubs *k* and *t*. In order to represent the effect of economies of scale on hub-connected flight, c_{iktj} is calculated by $\beta \cdot c_{ik} + \alpha \cdot c_{kt} + \beta \cdot c_{tj}$. $\beta \cdot c_{ik}$ and $\beta \cdot c_{tj}$ are used to approximate the transportation costs between a non-hub node and a hub. $\alpha \cdot c_{kt}$ represents the transportation cost on inter-hub link.

f_k : the fixed cost to set up hub *k*.

U_k : the upper bound of capacity at hub *k* for inbound, outbound, and transshipping flows.

U_k^T : the upper bound of capacity at hub *k* for transshipping flows only.

V : a very large positive number.

Z : the objective value representing the total costs.

The variables x_{ij} and x_{iktj} are capable of representing all different types of services and paths. In addition, x_{iktj} is able to differentiate the segments between hubs and/or non-hubs. This also enables the calculation of different discounts on transportation costs for different types of segments. The definitions of x_{ij} and x_{iktj}, and the corresponding services, paths, and cost discounts are shown in Table 1.

The economies of scale in transportation costs can be represented as a concave function. This study uses discount factors α and β to approximate the concave functions. The assignment problem of optimally allocating hub locations is already NP-Hard. In addition, this study considers optimal service network design and economies of scale effect in transportation costs which tremendously increase the complexity of the problem. Unless one can develop an efficient solution procedure; otherwise, using non-linear concave functions to model the transportation costs will not be practical because the solution would be very difficult to obtain even for a small problem. Though the literature has studied the hub location problem with concave cost functions and pointed out the advantages of using concave cost functions (O'Kelly and Bryan, 1998), their solution method is to use piece-wise linear functions to approximate the non-linear concave functions (as shown in Figure 1a) and avoids directly dealing with integer problems with non-linear objective function. By identifying the relationship between discount factors and flows, the discount factors α and β can also effectively approximate the flow-dependent costs, as shown in Figure 1b. If one can collect sufficient data from the field and identify the flow-dependent relationship of discount factors, the proposed model can be modified to account for the economies of scale in transportation costs. The use of discount factors reduces non-linear cost functions to linear cost functions, while maintains the flow-dependent effects. The problem, hence, becomes tractable and is able to be solved by existing solution algorithms or optimization software.

TABLE 1: The definition of x_{ij} and x_{iktj}

Variable	Service	Path	Cost discount
x_{ij}	non-stop	ⓘ → ⓙ non-hub non-hub	no
x_{kktt}	non-stop	ⓚ → ⓣ hub hub	α
x_{ikkk}	non-stop	ⓘ → ⓚ non-hub hub	β
x_{kkkj}	non-stop	ⓚ → ⓙ hub non-hub	β
x_{ikkj}	one-hub-stop	ⓘ → ⓚ → ⓙ non-hub hub non-hub	$i \to k : \beta$ $k \to j : \beta$
x_{kktj}	one-hub-stop	ⓚ → ⓣ → ⓙ hub hub non-hub	$k \to t : \alpha$ $t \to j : \beta$
x_{iktt}	one-hub-stop	ⓘ → ⓚ → ⓣ non-hub hub hub	$i \to k : \beta$ $k \to t : \alpha$
x_{iktj}	two-hub-stop	ⓘ → ⓚ → ⓣ → ⓙ non-hub hub hub non-hub	$i \to k : \beta$ $k \to t : \alpha$ $t \to j : \beta$

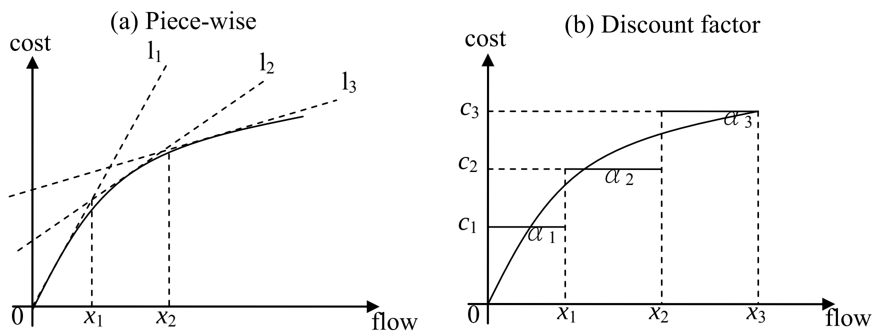

FIGURE 1: Approximation approaches to concave cost functions

Based on different hub functions, we proposed two different models. The first model addresses the capacitated hub location problem, where the hubs function to handle inbound, outbound, and transshipping flows. Therefore, the capacity limitation in the first model considers all traffic passing the hubs. The second model focuses on hubs designed mainly to handle transshipping flows, and the capacity constraint only accounts for the flows transshipping via the hubs. The first model (P1) can be formulated as follows:

Problem P1

Minimize $$z=\sum_{i\in N}\sum_{j\in N:j\neq i} d_{ij}c_{ij}x_{ij}+\sum_{i\in N}\sum_{j\in N:j\neq i}\sum_{k\in N}\sum_{t\in N} d_{ij}c_{iktj}x_{iktj}+\sum_{k\in N} f_k s_k \tag{1}$$

subject to

$$x_{ij}+\sum_{k\in N}\sum_{t\in N} x_{iktj}=1\,,\ \ \forall i,j\ :\ i\neq j \tag{2}$$

$$\sum_{i\in N} x_{ik}+\sum_{i\in N} x_{ki}\leq V(1-s_k)\,,\ \ \forall k:\ i\neq k \tag{3}$$

$$\sum_{t\in N} x_{kkti}+\sum_{t\in N} x_{itkk}\geq 2s_k\,,\ \ \forall i,k:\ i\neq k \tag{4}$$

$$x_{kktt}\geq s_k+s_t-1\,,\ \ \forall k,t:\ k\neq t \tag{5}$$

$$V\left(\sum_{i\in N}\sum_{j\in N}\sum_{t\in N}\left(x_{iktj}+x_{itkj}\right)-\sum_{i\in N}\sum_{j\in N} x_{ikkj}\right)\geq s_k\,,\ \ \forall k:\ i\neq j\,,\ i\neq k\,,\ j\neq k \tag{6}$$

$$\sum_{i\in N}\sum_{j\in N}\sum_{t\in N} d_{ij}\left(x_{iktj}+x_{itkj}\right)-\sum_{i\in N}\sum_{j\in N} d_{ij}x_{ikkj}\leq U_k s_k\,,\ \ \forall k:\ i\neq j \tag{7}$$

$$0\leq x_{ij}\leq 1\,,\ \ \forall i,j\ :\ i\neq j \tag{8}$$

$$0\leq x_{iktj}\leq 1\,,\ \ \forall i,k,t,j\ :\ i\neq j \tag{9}$$

$$s_k\in\{0,1\}\ ,\ \ \forall k \tag{10}$$

The objective function (1) minimizes the total cost and contains three terms. The first term is the transportation costs for nonstop services between non-hub nodes. The second term is the transportation costs for hub-connected services, including one-hub-stop and two-hub-stop services. The third term is the fixed cost of setting up hubs. Constraint (2) describes that the summation of the flow fractions between a pair of origin and destination is equal to one. This ensures conservation of flow between each origins-destination pair. Constraint (3) specifies that the non-stop flight between nodes k and i should be represented as x_{kkki} or x_{ikkk}, not x_{ik} or x_{ki}, when node k is a hub. Constraint (4) states that only non-stop and one-hub-stop services are allowed when either origin or destination is a hub. Constraint (5) only allows non-stop services when both origin and destination are hubs. According to assumption 7, the hub-and-spoke system enjoys different levels of economies of scale for both inter-hub traffic and traffic between non-hub nodes and hubs. This gives model (P1) an advantage for the construction of a hub-and-spoke network. However, in order to enjoy the economies of scale from the hub-and-spoke system, it might be possible to set up a hub that only provides non-stop services without any transshipping flows. Constraint (6) precludes this special case from happening in the proposed model. According to constraint (6), at least some of hub-stop routes passing through hub k have to be in service if node k is a hub. Constraint (7) is the capacity limitation on hubs which accounts for the inbound, outbound, and transshipping flows in problem (P1). Constraints (8) and (9) restrain the flow fractions within the range from 0 to 1. Constraint (10) requires the decision variable s_k to be either 0 or 1.

The second model formulates the problem where the function of hubs is handling transshipping traffic. Hence, the capacity limitation on hubs only takes into consideration transshipping flows. The second model (P2) can be formulated as follows:

Problem P2

$$\text{Minimize } z = \sum_{i\in N}\sum_{j\in N: j\neq i} d_{ij}c_{ij}x_{ij} + \sum_{i\in N}\sum_{j\in N: j\neq i}\sum_{k\in N}\sum_{t\in N} d_{ij}c_{iktj}x_{iktj} + \sum_{k\in N} f_k s_k \tag{11}$$

subject to

$$x_{ij} + \sum_{k\in N}\sum_{t\in N} x_{iktj} = 1, \quad \forall i,j : i\neq j \tag{12}$$

$$\sum_{i\in N} x_{ik} + \sum_{i\in N} x_{ki} \leq V(1-s_k), \quad \forall k : i\neq k \tag{13}$$

$$\sum_{t\in N} x_{kkti} + \sum_{t\in N} x_{itkk} \geq 2s_k, \quad \forall i,k : i\neq k \tag{14}$$

$$x_{kktt} \geq s_k + s_t - 1, \quad \forall k,t : k\neq t \tag{15}$$

$$V\left(\sum_{i\in N}\sum_{j\in N}\sum_{t\in N}\left(x_{iktj} + x_{itkj}\right) - \sum_{i\in N}\sum_{j\in N} x_{ikkj}\right) \geq s_k, \quad \forall k : i\neq j,\ i\neq k,\ j\neq k \tag{16}$$

$$\sum_{i\in N}\sum_{j\in N}\sum_{t\in N}\left(x_{iktj} + x_{itkj}\right) - \sum_{i\in N}\sum_{j\in N} x_{ikkj} \leq Vs_k, \quad \forall k : i\neq j \tag{17}$$

$$\sum_{i\in N}\sum_{j\in N}\sum_{t\in N} d_{ij}\left(x_{iktj} + x_{itkj}\right) - \sum_{i\in N}\sum_{j\in N} d_{ij}x_{ikkj} \leq U_k^T s_k, \quad \forall k : i\neq j, i\neq k, j\neq k \tag{18}$$

$$0 \leq x_{ij} \leq 1, \quad \forall i,j : i\neq j \tag{19}$$

$$0 \leq x_{iktj} \leq 1, \quad \forall i,k,t,j : i\neq j \tag{20}$$

$$s_k \in \{0,1\}, \quad \forall k \tag{21}$$

The objective function (11) of the model (P2) is the same as the objective function (1). Therefore, it also minimizes the total cost, which contains transportation costs and hub setup costs. Constraints (12) to (16) are the same as constraints (2) to (6) in problem (P1). Constraint (17) states that the hub-connected services via hub k can be used only when node k is a hub. If node k is not a hub, the hub-connected segments through hub k have to be disabled. Constraint (18) is the hub capacity constraint. Since the hub in problem (P2) mainly functions to transship flows, the flows are only calculated for transshipping traffic. Constraints (19) and (20) define the range of decision variables, and constraint (21) is an integrality constraint.

The proposed models (P1) and (P2) seek to design an optimal network which could be a point-to-point service, hub-and-spoke system, portions of point-to-point operations and portions of hub-and-spoke system, or even coexistence of point-to-point and hub-connected services in some flight routes. The network structure is not limited to point-to-point service or hub-and-spoke system. The models maintain the most flexibility, enabling designing of a network for all possible operations.

Problems (P1) and (P2) are both mixed integer programs. Suppose that we consider a network with n nodes. The number of decision variables is estimated in Table 2 and the number of constraints is listed in Table 3.

TABLE 2: The number of decision variables

Decision variables	Number of variables
s_k (integer)	n
x_{ij} (real)	n^2
x_{iktj} (real)	n^4
Total	$n^4 + n^2 + n$

TABLE 3: The number of constraints for (P1) and (P2)

Problem (P1)		Problem (P2)	
Constraints	Number of constraints	Constraints	Number of constraints
(2)	$n(n-1)$	(12)	$n(n-1)$
(3)	n	(13)	n
(4)	$n(n-1)$	(14)	$n(n-1)$
(5)	$n(n-1)$	(15)	$n(n-1)$
(6)	n	(16)	n
(7)	n	(17)	n
		(18)	n
Total	$3n^2$	Total	$3n^2 + n$

3. CASE STUDY AND SENSITIVITY ANALYSIS

The proposed models were tested using data on air freight flows between the top 10 cities in Taiwan and China ranked by the economic exchange activities between Taiwan and China (The Mainland Affairs Council, 2000; General Administration of Civil Aviation of China, 2000). Currently the air space between Taiwan and China is restricted due to political issues, and direct flights between Taiwan and China are not allowed. The 10 cities included in Table 4 accounted for 88% of the surveyed economic exchange in 1999. They should be sufficient to represent the air freight market after the restrictions on direct flights are relaxed. The geographical locations of the 10 cities are shown in Figure 2.

FIGURE 2: The geographical locations of the 10 cities

The data used to estimate the transportation and hub setup costs were collected from a leading airline company in Taiwan. Since some information is regarded as confidential

TABLE 4: The 10 cities and their demands

Node no.		1	2	3	4	5	6	7	8	9	10
	City	Taipei	Beijing	Tianjin	Shanghai	Nanjing	Hangzhou	Xiamen	Fuzhou	Gangzhou	Shenzhen
1	Taipei	0	8773	400	45649	2418	4672	10483	5912	39253	11941
2	Beijing	8773	0	100	54303	4857	4594	6349	3221	39438	16350
3	Tianjin	400	100	0	3673	100	100	100	100	3323	917
4	Shanghai	45649	54303	3673	0	100	100	7573	5301	36541	15969
5	Nanjing	2418	4857	100	100	0	100	1230	1063	6020	2532
6	Hangzhou	4672	4594	100	100	100	0	1533	790	18356	2414
7	Xiamen	10483	6349	100	7573	1230	1533	0	100	4432	1964
8	Fuzhou	5912	3221	100	5301	1063	790	100	0	2933	1021
9	Gangzhou	39253	39438	3323	36541	6020	18356	4432	2933	0	100
10	Shenzhen	11941	16350	917	15969	2532	2414	1964	1021	100	0

by the airline, the estimations might not be precise. However, they are sufficient to demonstrate the performance of the proposed models, and the conclusions from the testing results could still be valid in most applications. Previous studies suggest that a reasonable value of α is between 0.6 to 0.8 and a reasonable value of β is between 0.7 to 0.9 (O'Kelly, 1987; O'Kelly et al., 1996; Aykin, 1995). Two combinations of discount factors, ($\alpha = 0.6$, $\beta = 0.8$) and ($\alpha = 0.7$, $\beta = 0.9$), were used to test the proposed models. Detailed testing data are listed in Table 5.

TABLE 5: Testing data settings

Hub setup cost	Unit transportation cost	Discount factors
420,000,000	8.77	$\alpha = 0.6$, $\beta = 0.8$
(New Taiwan Dollars, NTD)	(NTD/km·ton)	$\alpha = 0.7$, $\beta = 0.9$

The modeling software GAMS in conjunction with the solver OSL were used to solve the problems. The test platform is a personal computer (Intel Pentium 4 CPU 3.40 GHz with 2.0GB RAM) with a Microsoft Windows XP operating system.

3.1 Sensitivity analysis for airport capacities

This section aims to analyze the influence of airport capacity limitations on hub locations and network structures. The capacity limitations for all airports varied between 10,000 tons and 400,000 tons. Figure 3 shows the objective values (in New Taiwan Dollars) of model (P1) with respect to changes in capacity limitations, and hub locations are displayed accordingly. The objective values are reduced as capacity limitations increase. This is because (1) the hubs are not located at the best sites due to insufficient capacity; (2) the demand cannot be transported through the least-cost path as a whole due to capacity limitations. Portions of the flows need to be routed through higher-cost paths. When capacity is less than 160,000 tons, none of the airports have enough capacity to handle transshipping flows. Therefore, all freight is transported through non-stop services. As the capacity increases, the model (P1) starts to select some airports as hubs in order to reduce transportation costs. In the case of ($\alpha = 0.6$, $\beta = 0.8$), the model tends to set up more hubs, because hub-connected services have lower transportation costs due to the lower discount factors in comparing with the case of ($\alpha = 0.7$, $\beta = 0.9$). Hence, the case of ($\alpha = 0.6$, $\beta = 0.8$) is more sensitive to changes of capacity limitations as opposed to the case of ($\alpha = 0.7$, $\beta = 0.9$). The hub locations and network structures change frequently for the combination of discount factors ($\alpha = 0.6$, $\beta = 0.8$). Since the case of ($\alpha = 0.7$, $\beta = 0.9$) is less sensitive to capacity limitation, no hubs are chosen until the capacity increases to 290,000 tons.

The similar tests were applied to model (P2) and the result is shown in Figure 4. The capacity limitations for all airports varied between 1,000 and 100,000 tons. The capacity constraint in model (P2) only accounts for transshipping flows. The result shows that problem (P2) is less sensitive to changes in capacity limitation than model (P1). The hub locations remain the same as capacity limitations change (see Figure 4).

Then we varied only the capacity limitation for a single airport. When there are no capacity limitations for all airports, models (P1) and (P2) choose nodes 2, 4, and 9 as hubs. Therefore, we only changed the capacity limitation for nodes 2, 4, and 9 respectively to see if the hub location and network structure will be different. Figures 5, 6, and 7 show the testing results for model (P1) and Figures 8, 9, and 10 for model (P2).

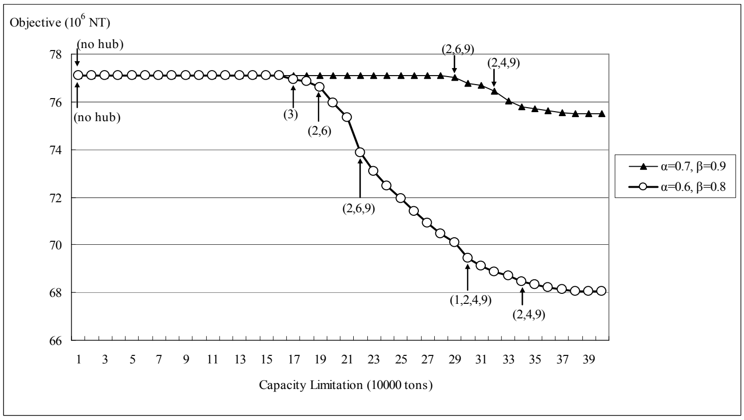

FIGURE 3: Model (P1) objective values versus capacity limitations of all airports under different discount factor combinations

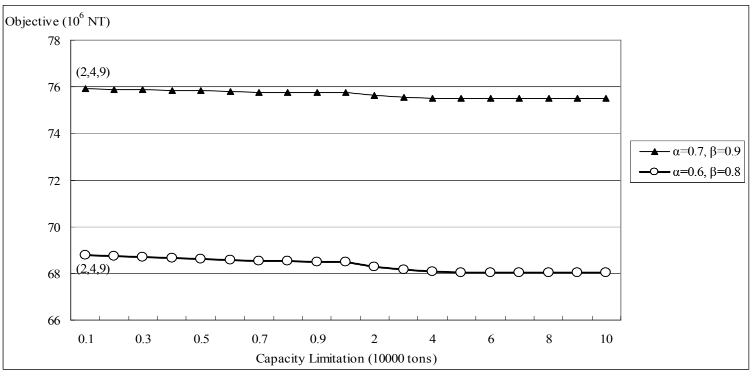

FIGURE 4: Model (P2) objective values versus capacity limitations of all airports under different discount factor combinations

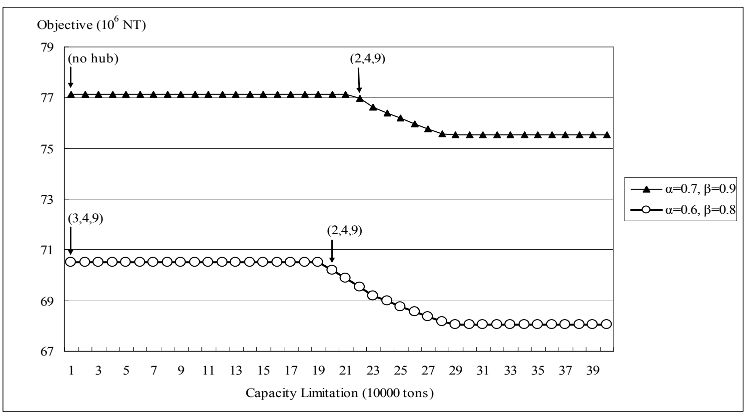

FIGURE 5: Model (P1) objective values versus capacity limitation at node 2 for different discount factor combinations

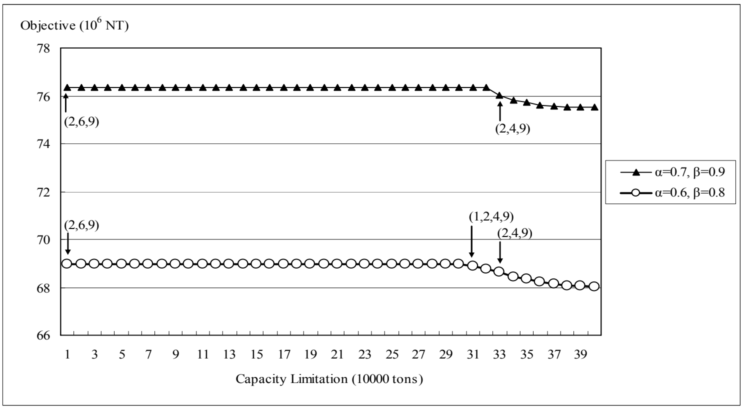

FIGURE 6: Model (P1) objective values versus capacity limitation at node 4 for different discount factor combinations

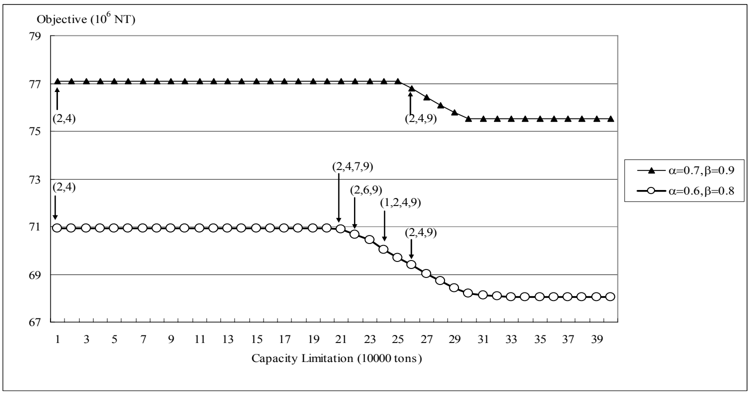

FIGURE 7: Model (P1) objective values versus capacity limitation at node 9 for different discount factor combinations

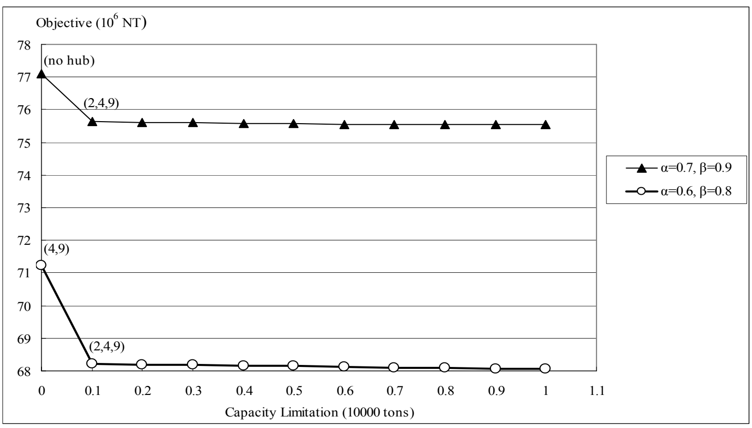

FIGURE 8: Model (P2) objective values versus capacity limitation at node 2 for different discount factor combinations

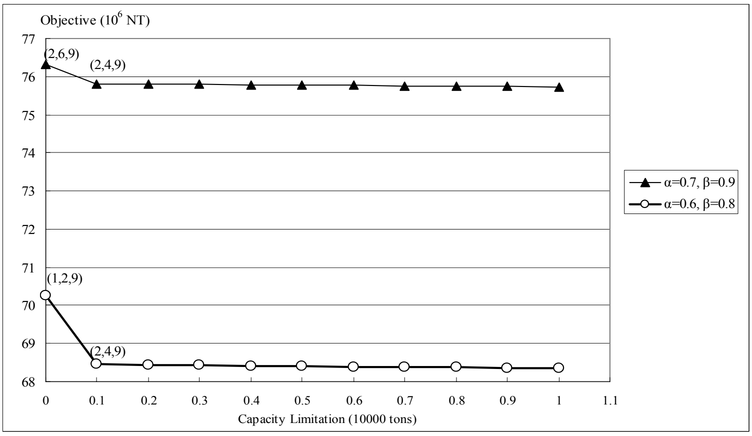

FIGURE 9: Model (P2) objective values versus capacity limitation at node 4 for different discount factor combinations

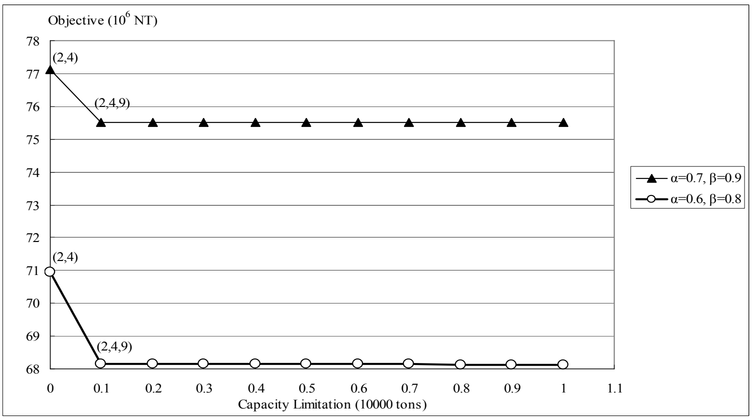

FIGURE 10: Model (P2) objective values versus capacity limitation at node 9 for different discount factor combinations

The capacity of node 2 increased from 10,000 to 400,000 tons, while keeping other airports at sufficient capacities. As shown in Fig. 5, in the case of ($\alpha = 0.6$, $\beta = 0.8$), model (P1) chose nodes 3, 4, and 9 as hubs when the capacity of node 2 is less than 200,000 tons, because node 2 does not have enough capacity to be a hub (as shown in Figure 5). For the case of ($\alpha = 0.7$, $\beta = 0.9$), model (P1) used point-to-point system to operate, as the capacity of node 2 is less than 210,000 tons (as shown in Figure 5). The model started to adopt a hub-and-spoke system after the capacity of node 2 was increased to 210,000 tons. As shown in the results, node 2 has the third highest demand and most of the hub-connected flights transship at node 2. This makes node 2 the most important hub in the network. Therefore, as the capacity of node 2 is insufficient and it could not operate as a hub, point-to-point operation became a better system. Figure 6 illustrates the testing results for the change of capacity of node 4 in model (P1). As the capacity of node 4 is short, hubs are located at nodes 2, 6, and 9. Since node 4 is a better hub location than node 6, model (P1) eventually chose node 4 as a hub when the capacity of node 4 became sufficient. For the case of ($\alpha = 0.6$, $\beta = 0.8$), as the capacity of node 4 is between 310,000 and 320,000 tons, node 4 becomes a hub, but it does not have enough capacity to transport all the transshipped flows. Therefore, the system requires one more hub, node 1, to transport the flows. Figure 7 shows the testing results for the change of capacity of node 9 in model (P1). The same reasoning from the above explanations can be applied to explain the change of hub locations in Figure 7.

The similar tests were applied to model (P2) and the results are shown in Figures 8, 9, and 10. The capacities for nodes 2, 4 and 9 were varied from 0 to 10000 tons respectively. We obtained the similar conclusion as the previous tests. The model (P2), whose capacity constraints only account for transshipped flows, is less sensitive to the change of capacity limitation comparing with model (P1). When a single hub's capacity was reduced, model (P2) would either close the hub (as shown in Figures 8 and 10) or choose another location instead (as shown in Figure 9). Unlike the tests in model (P1), the hub locations vary frequently as the hub capacity changes.

3.2 Sensitivity analysis for transportation costs

This section aims to analyze the influence of different transportation costs. In order to isolate the effect of change in transportation cost, airport capacities were set to be large enough to handle every situation. Since the difference between problems (P1) and (P2) is the capacity constraint and airport capacities have no effect in this analysis, the testing results based on (P1) and (P2) would be the same. Hence, in the following illustrations, we only show one of them.

Five different unit transportation costs were collected from the major airline companies in Taiwan and China to test the proposed models. As shown in Table 6, the hub location and network structure differ when the unit transportation cost changes. As the unit cost increases, the network is inclined to use more hub-stop flights and, hence, assigns more hubs. When the unit costs are 8.77 and 8.85, the number of hub-stop flights is 52. As the unit costs increase to 11.90, 13.53, and 14.73, the number of hub-stop flights increases to 70 and the number of hubs also increases to 4. This is because the hub-stop flights enjoy economies of scale through discount factors. Therefore, the flows tend to concentrate on hub-stop routes and reduce transportation costs by taking the discount.

TABLE 6: The comparison of different unit transportation costs

Unit transportation cost (NTD/km·ton)	Hub location	Non-stop flights	Hub-stop flights	Objective value
8.77	Beijing, Shanghai, Guangzhou	38	52	7551941692.89
8.85	Beijing, Shanghai, Guangzhou	38	52	7609336828.06
11.90	Taipei, Beijing, Shanghai, Gangzhou	20	70	9754240049.90
13.53	Taipei, Beijing, Shanghai, Guangzhou	20	70	10860207384.47
14.73	Taipei, Beijing, Shanghai, Guangzhou	20	70	11674416465.13

The previous analysis assumed that all the cities have the same unit transportation cost. However, the airline company may face different transportation costs in different cities. The transportation costs of outbound and inbound flights at three cities (Beijing, Shanghai, and Taipei) were increased to 1.25, 1.5, 1.75, and 2 times of the costs at other cities. The model chose Beijing, Shanghai, and Guangzhou as hubs when the unit transportation costs at all cities are the same. As the unit costs at Beijing, Shanghai, and Taipei increase, the model will give up Beijing and Shanghai and choose nearby cities of Tianjin and Hangzhou as hubs in order to avoid using the high-cost flights connecting to Beijing, Shanghai, and Taipei. As the unit transportation costs increase, the model tends to use more hub-stop flights to enjoy cost saving from discounts through hubs. The results are summarized on Table 7.

TABLE 7: The comparison of the increasing unit transportation cost at three cities

Cost increased ratio	Hub location	Non-stop flights	Hub-stop flights	Objective value
1	Beijing, Shanghai, Guangzhou	38	52	7551941692.89
1.25	Tianjin, Hangzhou , Guangzhou	14	76	8438518786.34
1.5	Tianjin, Hangzhou , Guangzhou	14	76	8992977495.25
1.75	Tianjin, Hangzhou , Fuzhou, Guangzhou	4	86	9412123586.29
2	Tianjin, Hangzhou , Fuzhou, Guangzhou	2	88	9763136013.67

4. SUMMARY AND CONCLUSIONS

This study addresses the capacitated airline network design problem. Air congestion has become a serious problem recently. The capacity limitations at airports cannot be ignored during the process of airline network planning. Hubs may be designed for different purposes according to different applications, and different functions of hubs require different airport resources. Therefore, two models were proposed to design the airline network based on different types of hubs. The first model is designed for the hubs where the inbound, outbound, and transshipping flows are equally important. The second model is used for the hubs that are primarily designed to handle the transshipping flows because the inbound and outbound flows are not significant.

The data from the air freight market between Taiwan and China was used to test the proposed models. The testing results suggested that the hub locations and network structures would be affected by different capacity constraints. Therefore, they are more sensitive to the model that accounts for inbound, outbound, and transshipping flows. The analysis showed that capacity limitations at airports are critical to airline network design and should be explicitly considered during the planning process. Different settings and combinations of transportation costs were also analyzed. The results showed that the use of hub-stop flights is encouraged in order to enjoy cost savings from economies of scale as the unit transportation cost increases. The models tended to locate hubs at cities with lower transportation costs where traffic could avoid taking flights with higher transportation costs. Analysis showed that a different cost structure may result in a totally different hub location and network structure.

The proposed models are path-based formulations which may become large mixed integer problems in practical application. Developing an efficient algorithm to solve the problems could be a direction for future work.

ACKNOWELEDGEMENTS

This material is based upon work supported by the National Science Council of Taiwan under Grant No. NSC 92-2211-E-415-001. The author would also like to thank the anonymous referees for their helpful comments and suggestions on the presentation of the paper.

REFERENCES

Aykin, T. (1994) Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem. European Journal of Operational Research, 79, 501-523.

Aykin, T. (1995a) Networking policies for hub-and-spoke systems with application to the air transportation system. Transportation Science, 29, 201-221.

Aykin, T. (1995b) The hub location and routing problem. European Journal of Operational Research, 83, 200-219.

Button, K. (2002) Debunking some common myths about airport hubs. Journal of Air Transport Management, 8, 177-188.

Campbell, J.F. (1994) Integer programming formulations of discrete hub location problems. European Journal of Operational Research, 72, 387-405.

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

General Administration of Civil Aviation of China (2000) Statistical Data on Civil Aviation of China 2000, China Civil Aviation Publisher. (in Chinese)

Klincewicz, J.G.. (1996) A dual algorithm for the uncapacitated hub location problem. Location Science, 4, 173-184.

O'Kelly, M.E. (1986a) The location of interacting hub facilities. Transportation Science, 20, 92-106.

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

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

O'Kelly, M.E. (1992) Hub facility location with fixed costs. Papers in Regional Science, 71, 293-306.

O'Kelly, M.E. and Bryan, D. (1998) Hub location with flow economies of scale. Transportation Research Part B, 32, 605-616.

O'Kelly, M.E. and Bryan, D. (2002) Interfacility interaction in models of hub-and-spoke networks. Journal of Regional Science, 42, 145-164.

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.

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

Skorin-Kapov, D., Skorin-Kapov, J. and O'Kelly, M.E. (1996) Tight linear programming relaxation of uncapacitated p-hub median problems. European Journal of Operational Research, 94, 582-593.

Sohn, J. and Park, S. (1998) Efficient solution procedure and reduced size formulations for p-hub location problems. European Journal of Operational Research, 108, 118-126.

The Mainland Affairs Council (2000) Cross-Strait Statistical Monthly. The Mainland Affairs Council, Executive Yuan, the Republic of China, Taiwan. (in Chinese)