A Meta-heuristic Approach for a hub-spoke air cargo network design problem

Sangheon Han

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 meta- 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.

Keywords: Meta-heuristic, hub network design, air cargo systems.

1. Introduction

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

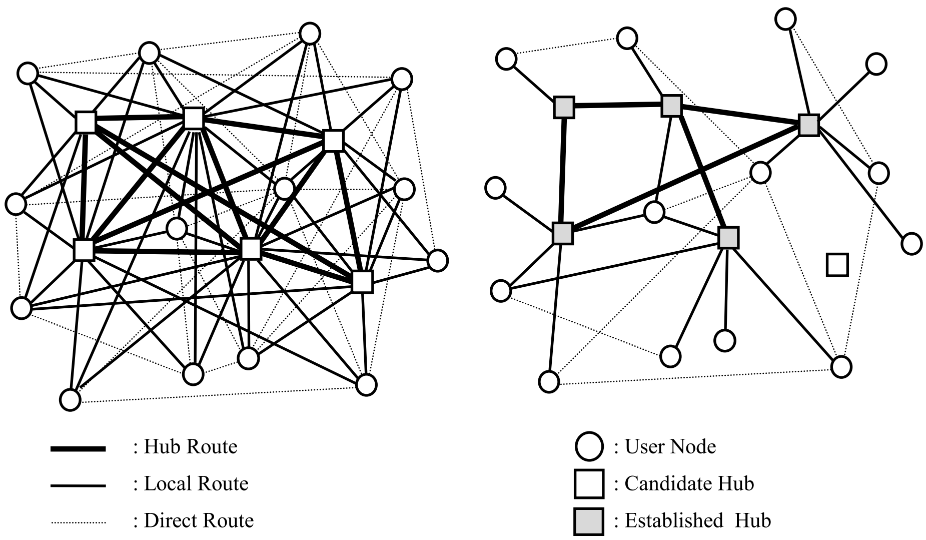
For air-cargo systems, HS structure has the same benefit as for passenger airlines. An example of an air-cargo system with HS structure is given in Figure 1, where user and hub nodes correspond to local and hub airports respectively; each arc represents a flight route. The demand at a user node is transported 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.

Our problem is specifically described as below.

(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.

(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 that 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



a) Given network configuration b) A feasible network design

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

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.

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.

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 0-1 variable denoting the demand of commodity *k* transported on arc (i, j),

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

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

$s.t.$

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

$$x_{ij}^k \leq y_{ij}, \qquad \{i,j\}\in E,\ k\in K, \tag{3}$$

$$x_{ji}^k \leq y_{ij}, \qquad \{i,j\}\in E,\ k\in K, \tag{4}$$

$$\sum_{i\in N} x_{ji}^k \leq z_j, \qquad j\in J,\ k\in K, \tag{5}$$

$$\sum_{i\in N}\sum_{j\in N} x_{ij}^k \leq h_k, \qquad k\in K, \tag{6}$$

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

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

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 to 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 they can be used to represent several quality constraints such as delivery time, damage and loss rates in cargo networks.

3. Genetic Algorithm Approach

The Genetic Algorithm (GA) is an adaptive heuristic search method based on population genetics. The basic concepts were developed by Holland (1975), while the practicality of using the GA to solve complex problems is demonstrated in Dejong (1975) and Goldberg (1989) . References and details about genetic algorithms can also be found for example in Alander (2000) and Mühlenbein (1997) respectively.

The creation of a new generation of individuals involves primarily four major steps or phases: representation, selection, recombination (crossover), and mutation. The representation of the solution space consists of encoding significant features of a solution as a chromosome, defining an individual member of a population. Typically pictured by a bit string, a chromosome is made up of a sequence of genes, which capture the basic characteristics of a solution. The recombination or reproduction process makes use of genes of selected parents to produce offspring that will from the next generation. It combines characteristics of chromosomes to potentially create offspring with better fitness. As for mutation, it consists of randomly modifying gene(s) of a single individual at a time to further explore the solution space and ensure, or preserve, genetic diversity. The occurrence of mutation is generally associated with low probability. A new population replaces those from the old one. A proper balance between genetic quality and diversity is therefore required within the population in order to support efficient search.

Although theoretical results that characterize the behaviour of the GA have been obtained for bit-string chromosomes, not all problems lend themselves easily to this representation. This is the case, in particular, for sequencing problems, like vehicle routing problem, where an integer representation is more often appropriate. We are aware of only one approach by Thangiah (1995) that uses bit string representation in vehicle routing context.

A basic scheme of a typical algorithm is as follows.

```
Randomly create an initial population
While not (termination condition)
   do
     Evaluate each member's fitness
     Kill the bottom x% elements of the population
       Let the fitness reproduce themselves
       Randomly select two members/parents (many other selection methods are also used)
       Perform crossover on the selected elements to generate two children
                                                   (many variations of crossover exist)
       Perform mutation
Endwhile
```

The problem **(P)** is a 0-1 Integer Problem (IP), which contains a hub location problem as well as a network design problem with hop-count constraints. Moreover, this problem is an uncapacitated network design problem having NP-hard computational complexity. 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 proposed a Genetic Algorithm in meta-heuristic methods to solve our comprehensive problem with ease. The Meta-heuristic method is very effective for this problem as it is for a combinatorial optimization problem (COP).

In this paper, we propose a coding method that is a matrix-form coding method. 2by2-method is very efficient to control lethal gene in the GA process. We set GA-parameters as below.

Selection Method	: Roulette wheel method
Population size	: 200
Crossover rate	: 0.3
Mutation rare	: 0.02
Generation	: 1,500

4. Computational Experiments

The solution procedure was coded in C, and test runs were performed on a PC (Intel Core 2 Duo /3GHz) 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 12 randomly generated problems by using our meta-heuristic. (M. Yoon & S. Han 1996)

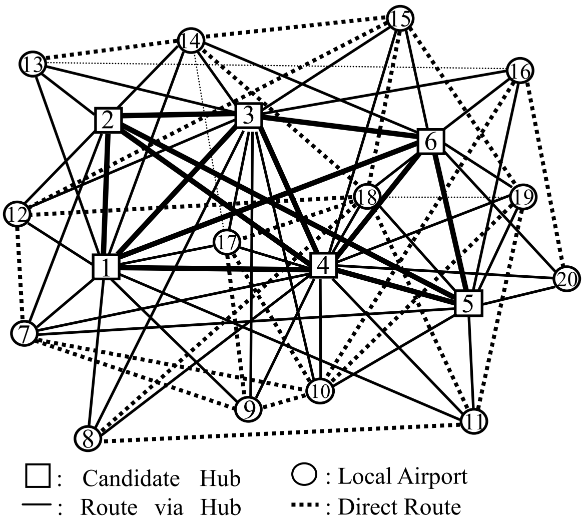


Figure 2. An example Network

To investigate the variations of hub-and-spoke networks according to the hop-count limit, we consider two types of hop-count constraints for the sample network. One has 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.

Table 1. Input parameters and computational results

g_j	f	h_k	Objective Value	No. of hubs[a]	No. of arcs				No. of routing paths			Computation Time (Sec.)
					Hub	Spoke	Direct	Total	1 Hub	2 Hubs	Non-Hub	
[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 – 7,000]	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		6,047,202	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

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 that being included in the network increases 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 half of the commodities are restricted to have 2-hops limit, the number of hub airports being established and the number of flight-routes required to transport commodities become greater than that of having 3-hops limit for all commodities, and the objective value is increased by 300 times.

To test the meta-heuristic method for more general cases, we generated the test problems randomly, 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 12 randomly generated problems with a hop-count limit of 3, and tried to solve them by using CPLEX program and meta- heuristic.

From previous research (M. Yoon & S. Han, 1996), we can find that 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, in 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 the 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.

Table 2. Randomly generated problems and computational results

Problem Set	|I|x|J|x|E|x|K|	[*a, b*]	*f*	Objective Value	No. of hubs	No. of routing paths	Computation Time (Sec.)
				GA	GA	GA	GA
III-1	10x40x350x780	[1,000 5,000]	10	165,978	4	72	143.2
III-2			50	179,997	3	59	190.1
III-3			100	199,452	3	53	159.2
III-4		[5,000 10,000]	10	193,214	3	64	198.8
III-5			50	189,774	2	44	135.6
III-6			100	201,201	2	43	108.4
IV-1	10x50x600x1225	[1,000 5,000]	10	260,807	4	79	491.1
IV-2			50	272,323	4	72	959.7
IV-3			100	287,201	3	57	98.5
IV-4		[5,000 10,000]	10	298,201	3	65	93.3
IV-5			50	261,789	3	63	160.6
IV-6			100	274,016	3	60	189.2

GA : the best primal objective value obtained by meta-heuristic (genetic algorithm)
a : the number of established hubs. b : the number of established arcs

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). The reader's attention is again called to the fact that our problem is so complex to include as subproblems in an integrated framework 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 good feasible solution for a large-sized network within a few minutes. The solution times are increased as the fixed costs of hub establishment and the fixed costs on arcs are increased. However, using meta-heuristics, most of the 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 the meta-heuristic method.

5. Conclusion

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, an 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 meta-heuristic.

Our model was tested on a sample network and randomly generated networks of varying cost structures. Eventhough 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 that 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.

References

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

[2] Dejong K.A. 1975. Analysis of the behavior of a classic of genetic adaptive system, PhD. thesis, Department of Computer and Communication Sciences, University of Michigan.

[3] Goldberg D.E. 1989. Genetic Algorithm in Search, Optimization, and Machine Learning, Addison-Wesley.

[4] Holland J.H., 1975. Adaption in Natural and Artificial Systems. University of Michigan Press, Ann Arbor.

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

[6] M. Yoon & S. Han, 2006,A dual based heuristic for a hub-spoke air cargo network design problem with hop-count constraints, *Journal of Scientiae Mathematicae Japonicae* 64-1, 77-90

[7] 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.

[8] Thangiah Sam 1995. Vehicle routing with time windows using genetic algorithms. In Application Handbook of Genetic Algorithms, New Frontiers, CRC Press, 2, 253-277.