Hindawi
Journal of Advanced Transportation
Volume 2019, Article ID 1812543, 14 pages
https://doi.org/10.1155/2019/1812543

Research Article

An Integrated Problem of p-Hub Location and Revenue Management with Multiple Capacity Levels under Disruptions

Yan-Ting Hou,[1,2] **Jia-Zhen Huo**,[1] **and Feng Chu**[2,3]

[1] *School of Economics and Management, Tongji University, Shanghai 200092, China*
[2] *Laboratory IBISC, Univ Evry, Université Paris-Saclay, Evry 91025, France*
[3] *School of Traffic and Transportation Engineering, Hefei University of Technology, Hefei 230009, China*

Correspondence should be addressed to Jia-Zhen Huo; huojiazhen@163.com

Received 20 October 2018; Accepted 17 February 2019; Published 1 April 2019

Academic Editor: Zhi-Chun Li

Copyright © 2019 Yan-Ting Hou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper considers an integrated hub location and revenue management problem in which a set of capacities is available from which one can be chosen for each hub and the disruption is considered in a star-star shaped airline network. We propose a two-stage stochastic programming model to maximize the profit of the network in which the cost of installing the hubs at different levels of capacities, the transportation cost, and the revenue obtained by selling airline tickets are considered. To provide flexible solutions, a hybrid two-stage stochastic programming-robust optimization model is developed by putting relative emphasis on a weighted sum of profit maximization. Furthermore, a sample average approximation approach is used for solving the stochastic programming formulation and a genetic algorithm approach is applied for both formulations. Numerical experiments are conducted to verify the mathematical formulations and compare the performance of the used approaches.

1. Introduction

For the development of the world's airlines in 2017, strong demand for air passenger service drove the expansion of the airline network exceeding 20000 unique city pairs as well as 9% net profit growth in excess of $38.0 billion [1]. However, the growth of the capacity of the infrastructure failed to keep pace at a similar growth rate as air traffic demand continued to soar [2]. Therefore, scarce capacity in the airline network has become a major issue in the aviation industry.

The integration of hub location and revenue management could be a potential solution by effectively building and using the capacity to achieve competitive profit. In the present status of relevant research, hub location problem lies at the heart of network design planning in transportation systems [3]. This research area focuses on locating the hubs from a set of nodes, activating a set of links, and routing the capacities of the network while optimizing a cost-based objective function. In fact, all these capacities correspond to all the seats on the flights that are operated in this network. In revenue management, airline tickets are considered instead of the seats. Revenue management is a backbone of the airline business by controlling the passenger demand for capacitated and perishable tickets with multiple fare classes to maximize the revenue [4]. Capacity control, the core of revenue management, is for allocating the seats to different fare classes before the flight departs with the objective of maximizing the revenue [5]. This means that the installed capacities can be well allocated by a revenue management problem. The profit calculated can reoptimize the installed network by solving the hub network problem. Therefore, hub location problem and revenue management problem are closely related.

Building sustainable, flexible, and affordable infrastructure is also the key to resolve the crisis of inadequate capacity in the network. Reference [6] addresses a capacitated single allocation hub location problem and determines the installed capacities for each hub from a finite set of capacity levels with different set-up costs. This strategy can largely improve the utilization of the infrastructure and reduce traffic delays. Besides, the disruptions, from low-level disobedience of crew instruction to major disruptions involving aircraft

diversion, passenger deplanement, equipment failure, jet fuel supply reliability, digital disruption, and adverse weather, could degrade the capacity [1, 7]. For example, airport authorities and regulators impose the restriction of the quota of available airport departure and arrival slots in adverse weather [8]. This could result in flight delay and profitability reduction. Hence, the impact of disruptions should be considered.

The uncertainty is another important factor to be considered while building the capacity. Stochastic programming is a widely used modeling paradigm to deal with uncertainty. However, the probability distribution function of the underlying stochastic parameters is hard to obtain. Robust optimization can be an effective alternative by computing over a convex uncertainty set of the parameters and bounding the maximum allowable deviation of the parameters from their nominal values to realize the worst-case optimization [9]. However, its result is more conservative. Therefore, [10] combines the benefits of both approaches and proposes a weighted sum of average and worst case profits to provide the flexibility by putting relative emphasis on average-case or worst-case value. Our paper also uses this idea to address uncertain demand.

Facing unbalanced growth between the capacity and the demand in the aviation industry, our research is motivated by the above solutions. Additionally, we present a two-stage stochastic programming formulation model to maximize the profit. In the first stage, the locations of p hubs, the capacity levels for the hubs, the connections between the hubs and the non-hubs, the protection levels of the tickets with different fare classes can be decided. The booking limits of the tickets with different fare classes can be obtained in the second stage. Moreover, we also provide a hybrid two-stage stochastic programming-robust optimization model to maximize a weighted sum of the profit. To solve the problem, a genetic algorithm (GA) approach is applied. For the stochastic programming model, we also use a sample average approximation approach. Numerical experiments demonstrate the effectiveness of these two formulations.

Our contribution is providing a novel insight of solving unbalanced growth between the demand and the capacity in the airline industry. We present a two-stage stochastic programming to describe the interaction between hub location and revenue management and provide a flexible strategy for the infrastructure by considering a set of capacities available for each hub under disruptions. In addition, a hybrid two-stage stochastic programming-robust optimization is investigated to maximize a weighted sum of the profit. Our solutions can well tackle this issue.

The rest of this paper is organized as follows. Section 2 discusses relevant literature. Section 3 presents the integrated hub location and revenue management problem with multiple capacity levels under disruptions. Section 4 gives two mathematical models. Section 5 explains the research approaches in detail. Section 6 provides numerical experiments and analyzes the results. Section 7 discusses the conclusions.

2. Literature Review

2.1. An Integrated Hub Location and Revenue Management Problem. Hub location problem is firstly proposed by [11]. This research focuses on the selection of a set of nodes to place hub facilities and the links of connecting the origins and the destinations via the hubs. Classical hub location problem mainly focuses on minimizing the cost under the assumption of satisfying all demands. However, from a profit point of view, some origin-destination ($O\&D$) pairs are not necessary to be served if they are not profitable. As such, the hub location problem by considering the trade-off between the revenue and the cost to maximize the profit has been discussed. The obtained revenue can be classified as (1) from captured flows, e.g., [12–15]; (2) from an integration of pricing and hub location, e.g., [16, 17]; and (3) from an integration of revenue management and hub location, e.g., [18, 19]. In this paper, we address an integrated problem of hub location and revenue management.

The research of revenue management is introduced by Littlewood in 1972. Revenue management evolves from five streams: pricing, capacity control, overbooking, auctions, and forecasting [20]. Our study is relevant to capacity control. Early publications provide effective analysis and the approaches of revenue management on single flight leg ([21, 22]). Progress in the revenue management problem has been moved forward to larger network since the research of network revenue management was proposed by [23]. This stream can be classified by the network topology: on the one hand, for multiple legs without a specific network structure, some research ([24, 25]) decomposes multiple legs into many single-leg problems and then solves each single-leg revenue management problem. The research on the hub-and-spoke network also has received significant attention ([26–28]) on the other hand. Reference [26] considers a network revenue management problem in a hub-and-spoke network consisting of a hub and N spokes. Reference [27] studies the capacity allocation in a two-airline alliance under competitions in a combined hub-and-spoke network. Reference [28] addresses an airline revenue management problem with consideration of a competitor's behavior under simultaneous price and capacity competitions in a hub-and-spoke network. Reference [29] explores the impact of the structural properties in a hub-to-hub network revenue management problem. As can be seen from the above literature, we conclude that the network structure plays an important role in airline revenue management. For reviews on revenue management, we refer the readers to, e.g., [30, 31].

Until [18], the integrated research of revenue management and hub location in the airline industry is proposed. In this research, they determine the booking limits for different fare classes and design an uncapacitated single allocation hub network with hub-stop and non-stop $O\&D$ pairs. They propose a two-stage stochastic programming formulation to maximize the profit, including the revenue from ticket sales and the cost of transportation and hub installation. In the first stage, the locations of the hubs and the central hub, the links between the hubs and the non-hubs, and the protection levels for different fare classes are determined. The booking limits

of the tickets should be decided depending on the protection levels in the second stage. They propose a hybrid optimization method by combining GA with caching technique and exact solution. Compared to pure GA, their results outperform in computational time and high quality on the instances with up to 20 nodes. Reference [19] continues the integrated hub location and revenue management problem in a star-star hub network. They formulate a two-stage stochastic programming to maximize the profit. They propose two hybrid optimization methods consisting of ordinary GA and modified caching hybrid GA. Their results show the efficiency of their methods on the instances with up to 25 nodes. In our paper, we also consider the integrated research of the hub location problem and the revenue management problem. The above studies have explored the interaction between the hub location problem and the revenue management problem, but the discussion only focuses on an uncapacitated hub network.

2.2. Hub Location with Multiple Capacity Choices. Hub location problem can be classified as capacitated and uncapacitated depending on whether the hub has limited capacity [32]. Complete review of hub location problem can be found in [3, 33]. Our paper focuses on a capacitated hub location problem (CHLP), which is firstly proposed in [34]. In their hub network, all traffic can be routed via either one or two hubs and the hubs are fully interconnected. The direct connection is not allowed between the two non-hubs. They present a mixed integer linear programming (MILP) formulation with fewer variables and constraints to minimize the cost of transportation and installing the hubs. Beyond that, a variant of CHLP in which a set of capacities is assumed to be available for each hub is discussed by [6].

Earlier studies aim at a CHLP with multiple capacity choices for single or multiple allocation case. The diversity between single allocation and multiple allocation is whether each non-hub is linked to a single hub or more than one hub. Regarding the capacitated single allocation hub location problem (CSAHLP), reference [35] considers a CSAHLP with multiple capacity choices and single commodity. Four MILP formulations for this problem are presented by introducing new variables and modifying the formulations based on [34, 36, 37]. To identify a model which can solve the problem optimally, they compare the lower bounds provided by the linear relaxations of these formulations. Reference [38] introduces the balancing requirements into the research [6]. The balancing requirements indicate that the difference between the maximum and the minimum number of the spokes connected to the hub should not overpass a maximum value. They propose two MILP formulations to minimize the setup cost and the routing cost. Their results show that the superior model can be confirmed in terms of the lower bounds provided by the linear relaxations of these two formulations for the instances up to 50 nodes. Later, dynamic hub location problem with multiple capacity levels is introduced. Reference [39] considers multiple periods in a multi-capacity single allocation hub location problem. In their assumptions, the hub can be closed or resized and the non-hub can become a hub in each period. They present four quadratic MILP formulations including two flow-based and two path-based models. The objective is to minimize the cost of the connection from the non-hub to the hub, the transportation from one hub to another hub, the distribution from the hub to the non-hub, opening the hub, resizing the hub, and closing the hub. Their results show that the path-based formulations outperform the flow-based formulations for almost all instances. Reference [40] addresses a multi-period capacitated hub location problem for both single allocation and multiple allocation cases. In each time period, open new hubs and expand the capacity of existing hubs are allowed. They use MILP formulation to minimize the cost of transportation, operation on hubs and hub links, capacity establishment, and hub installation over the planning horizon. They also enhance the model by proposing valid inequalities. Their results can be solvable in reasonable time for the instances with up to 64 nodes and confirm the importance of time dimension and the multi-period nature in hub network. A further step forward for the research on the multi-capacity hub location problem is to discuss the relevance with the level of congestion. Reference [41] considers a CHLP with service time and congestion for both single allocation and multiple allocation cases. They assume that service time includes travelling time through the network and handling time at hubs. The congestion of the hubs can lead to an increase in handling time. The direct connection is allowed between the two non-hubs. The MILP formulation with service time and congestion is used to minimize the cost of transportation and installing the hubs with different capacities. They also model the direct connection between the two non-hubs in multiple allocation case. They analyze the impact of different factors including service time limit, congestion factor, fixed cost, and the capacity. The above literature only discusses on building a multi-capacity hub network from a cost perspective. Their assumption is that all the nodes and the links should be served. However, if an $O\&D$ pair cannot generate a good profit, then there is not enough incentives to serve this $O\&D$ pair.

To the best of our knowledge, our paper fills gap in existing literature and advances the research in addressing the issue of unbalanced development between the capacity and the demand in the airline industry by providing competitive profit, giving flexible infrastructure and reducing the impact of disruptions. The notation and the formulation of our problem are presented in the next section.

3. Problem Statement

In this section, we start by presenting our problem in the first subsection. We use the network terminology in [42]. Afterwards, the assumptions and the parameters are depicted in the remaining subsections, respectively. Our problem is inspired by the integrated research of hub location and revenue management developed by [19], the research on the hubs with multiple capacity levels based on [41], and the reseach on the disruptions considered by [7].

3.1. Problem Definition. In a star-star network, the set I includes the nodes of both non-hubs and potential hubs. We

denote the non-hub as i and the hub as j. A central hub 0 is given. In one star topology, the center is the central hub 0 and the spokes are the hubs. The connections are between the hubs and the central hub 0. We name these connections as the primary (trunk) links in this star network. The other star topology is an implementation of a hub-to-non-hub distribution paradigm. This connection can be defined as the secondary (spoke) link. In our problem, any $O\&D$ pair from the origin i to the destination m can pass through either one hub or two hubs. Figure 1 illustrates this star-star hub network in our problem.

Given that the fixed installation cost of the central hub 0 is F'. Once the hub j is located, only one level q with capacity Γ^q should be chosen from a set W_j containing multiple capacity levels for hub j. And a fixed cost F_j^q is incurred accordingly. The maximum capacity on the primary link between the hub j and the central hub 0 is No_{j0}. And the maximum capacity on the secondary link from the non-hub i to the hub j is limited by No_{ij}. Transportation costs on the primary and secondary links are represented as C_{1k} and C_{2k}, respectively. But C_{1k} is lower than C_{2k} because the economies of scale play a role in reduced unit cost on the primary link between the hub and the central hub, where $C_{1k} = \varphi C_{2k}$, φ is a discount factor and $\varphi \in [0, 1]$. In our paper, the number of the flights represents the flow in this network, which is relevant to (1) the aircraft type operated on different types of the links. Only the aircraft type with capacity Q_1 is available on the primary link and the aircraft type with capacity Q_2 is for the secondary link. (2) the protection level of the tickets.

Given that there are K fare classes for each flight. The protection level of the tickets and the demand for fare class k impact the booking limit of airline ticket. Once the protection level is set, the impact of future adjustments can be ignored. We consider that the uncertainties are represented by stochastic parameters. The uncertainty can be captured by a finite set S of scenarios. Each scenario s occurs with a probability p_{imk}^s, $\sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\sum_{s\in S} p_{imk}^s = 1$. Each scenario s is relevant to the demand and disruption variation. The uncertain demand of airline tickets for the fare class k under scenario s is d_{imk}^s. The probability of the disruption at hub j under scenario s is α_j^s.

In our problem, a deterministic model is difficult to describe the changes in the demand. As such, we present a two-stage stochastic programming for our problem. The first-stage decision concerns (1) the locations of the hubs as well as their capacity levels, (2) the connections between the nodes on both the primary and secondary links, and (3) the protection levels of the tickets with different fare classes. In the second stage, we seek the decision on the booking limits of airline tickets. The objective is to maximize the profit, which contains several components: (1) the revenue from tickets sales, (2) the installation cost for the central hub and for the hubs with specific capacity levels, and (3) the transportation cost on both primary and secondary links. Taking the profit into consideration, our problem measures the trade-off between the revenue by selling airline tickets and the cost by building a capacitated hub network with multiple hub capacity levels.

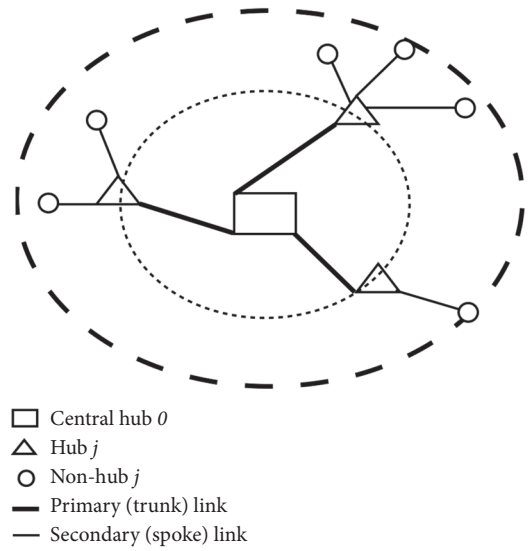


Figure 1: Airline star-star network.

3.2. Assumptions. We start itemizing the following assumptions:

(i) The number of the hubs is predefined as P.

(ii) The location of the central hub 0 is predefined.

(iii) Each non-hub i is connected to at most one hub j, but direct connection between the two non-hubs is not allowed. Similarly, each hub j should connect to the central hub 0 without direct connection between the two hubs.

(iv) A hub j can process its own outbound flow and distribute its own inbound flow.

(v) Hub j is capacitated. We assume that the capacity of hub j indicates that the maximum flow can be processed.

(vi) All the flows traverse at least one hub and at most two hubs (not including the central hub).

(vii) Only one capacity level is installed at each hub j. The installation cost is relevant to the capacity level that has been chosen for the hub.

(viii) The uncertain demand for each fare class is independent.

(ix) The hub is assumed to be at risk of the disruptions, except for the central hub 0.

(x) The arrival sequence of the demand is not considered.

(xi) Passengers who buy the tickets must appear at the time of departure. Cancellations and no-shows are not considered.

3.3. Parameters and Decision Variable. Before presenting our model, we continue to use the notations in [19] as follows:

Indices.

i, m: indices for node.

j: index for hub.

s: index for scenario.

k: index for customer class.

q: index for capacity level.

Sets.

I: set of nodes.

P: the number of the hubs needs to be built.

S: set of scenarios.

K: set of fare classes.

W: set of capacity levels available for each hub.

Parameters.

f_{j0}: distance on the primary link from hub j to central hub 0, $j \in I$.

f_{ij}: distance on the secondary link from non-hub i to hub j, $i, j \in I$.

C_{1k}: transportation cost (per unit of flow and per unit of distance) on the primary link from hub j to central hub 0 at fare class k, $k \in K$.

C_{2k}: transportation cost (per unit of flow and per unit of distance) on the secondary link at fare class k, $k \in K$.

d^s_{imk}: uncertain demand per $O\&D$ pair from the origin i to the destination m at fare class k under scenario s, $s \in S$, $i, m \in I$ and $k \in K$.

p^s_{imk}: the probability of each scenario s for $O\&D$ pair between the origin i and the destination m at fare class k, $s \in S$, $i, m \in I$ and $k \in K$.

r_{imk}: ticket price on $O\&D$ pair from the origin i to the destination m at fare class k, $i, m \in I$ and $k \in K$.

No_{j0}: capacity on the primary link from hub j to central hub 0, $j \in I$.

No_{ij}: capacity on the secondary link from non-hub i to hub j, $i, j \in I$.

Q_1: the largest seating capacity available for each aircraft on the primary link.

Q_2: the largest seating capacity of each aircraft on the secondary link.

Γ^q: capacity of a hub at capacity level q, $q \in W$.

F^q_j: installation cost of hub j at capacity level q, $j \in I$ and $q \in W$.

F': installation cost of central hub 0.

M: a very large integer.

α^s_j: the probability of disruption at hub j under scenario s, $\alpha^s_j \in [0, 1]$, $j \in I$ and $s \in S$.

Variables.

y^s_{imk}: booking limit of airline tickets for each fare class k on $O\&D$ pair from the origin i to the destination m under scenario s.

Z_{imk}: protection level of the tickets for each fare class k on $O\&D$ pair from the origin i to the destination m.

x_{jj}: binary decision variable. x_{jj} equals to 1 if node j becomes a hub node and 0 otherwise.

x_{ij}: binary decision variable. If non-hub i is routed to hub j, x_{ij} equals to 1 and 0 otherwise.

v^q_j: binary decision variable. v^q_j equals to 1 if capacity level q is chosen for hub j and 0 otherwise.

With these sets of variables and parameters, we can obtain the mathematical models in the next section.

4. Mathematical Formulations

In this section, we present two models of the integrated hub location and revenue management problem with considerations of multiple capacity levels and network disruptions: (1) a two-stage stochastic programming model in Section 4.1 and (2) a hybrid two-stage stochastic programming-robust optimization model in Section 4.2.

4.1. Stochastic Programming Model. In this section, we give nonlinear and linear formulations to describe our problem. First, a nonlinear model can be expressed as the following deterministic equivalent program [43]:

$$\begin{aligned}\max \quad & \sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\sum_{s\in S} p^s_{imk} r_{imk} y^s_{imk} - \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}\sum_{s\in S} C_{1k}\left(p^s_{imk} f_{j0}\left\lceil \frac{y^s_{imk}}{Q1}\right\rceil + p^s_{mik} f_{0j}\left\lceil \frac{y^s_{mik}}{Q1}\right\rceil\right)\left[x_{ij}\left(1-x_{mj}\right)\right] \\ & - \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}\sum_{s\in S} C_{2k}\left(p^s_{imk} f_{ij}\left\lceil \frac{y^s_{imk}}{Q2}\right\rceil + p^s_{mik} f_{ji}\left\lceil \frac{y^s_{mik}}{Q2}\right\rceil\right) x_{ij} - \sum_{j\in I}\sum_{q\in W} F^q_j x_{jj} - F'\end{aligned} \tag{1}$$

$$\text{Subject to} \quad \sum_{j\in I} x_{ij} \le 1 \quad i \in I \tag{2}$$

$$\sum_{j\in I} x_{jj} = P \tag{3}$$

$$x_{ij} \le x_{jj} \quad i, j \in I \tag{4}$$

$$y_{imk}^{s} \le d_{imk}^{s} \quad i, m \in I,\ k \in K,\ s \in S \tag{5}$$

$$y_{imk}^{s} \le Z_{imk} \quad i, m \in I,\ k \in K,\ s \in S \tag{6}$$

$$\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil \frac{Z_{imk}}{Q1}\right\rceil x_{ij}\left(1-x_{mj}\right)+\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil \frac{Z_{mik}}{Q1}\right\rceil x_{ij}\left(1-x_{mj}\right) \le No_{j0}x_{jj} \quad j \in I \tag{7}$$

$$\sum_{m\in I}\sum_{k\in K}\left\lceil \frac{Z_{imk}}{Q2}\right\rceil+\sum_{m\in I}\sum_{k\in K}\left\lceil \frac{Z_{mik}}{Q2}\right\rceil \le \sum_{j\in I}No_{ij}x_{ij}+Mx_{ii} \quad i \in I \tag{8}$$

$$\sum_{q\in W} v_j^q \le x_{jj} \quad j \in I \tag{9}$$

$$\begin{aligned}&\sum_{q\in W}\left(1-\alpha_j^s\right)\Gamma^q v_j^q\\ &\ge \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil \frac{Z_{imk}}{Q1}\right\rceil x_{ij}\left(1-x_{mj}\right)+\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil \frac{Z_{mik}}{Q1}\right\rceil x_{ij}\left(1-x_{mj}\right)+\sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\left\lceil \frac{Z_{imk}}{Q2}\right\rceil x_{ij}\\ &\quad+\sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\left\lceil \frac{Z_{mik}}{Q2}\right\rceil x_{ij} \quad j \in I,\ s \in S\end{aligned} \tag{10}$$

$$x_{ij},\ x_{jj},\ v_j^q \in \{0,1\} \quad i, j \in I,\ q \in W \tag{11}$$

$$y_{imk}^{s},\ Z_{imk} \in Z_{+} \quad i, m \in I,\ k \in K,\ s \in S \tag{12}$$

The objective function (1) is to maximize the profit. The profit is obtained by subtracting the cost from the revenue. In the first term, the revenue of airline tickets with different fare classes for all $O\&D$ pairs with the origins i and the ends m is computed. The second term denotes the transportation cost on the primary link between the hub and the central hub. The third term represents the transportation cost on the secondary link between the non-hub and the hub. The fourth term represents the installation cost of the hubs with the chosen capacity levels. The last term represents the installation cost of the central hub. Constraint (2) indicates that each non-hub must be allocated to at most one hub. That is, $\sum_{j\in I} x_{ij} = 1$ indicates that each non-hub allocates to one hub. $\sum_{j\in I} x_{ij} < 1$ indicates that not all P hubs need to be connected in the network when less profit incurs. The number of the hubs is P due to constraint (3). Constraint (4) guarantees that the non-hub only can be allocated to each installed hub. Constraints (5) and (6) explain that the booking limit of airline tickets at fare class k is related to uncertain demand and the protection level of airline tickets under scenario s. Constraints (7) and (8) restrict the total flows on the primary and secondary links. Constraint (7) restricts the total flow on sending and receiving directions of the links. The flow denotes the number of the flights with the aircraft type Q_1. This $O\&D$ pair has the link between the hub and the central hub only when $x_{ij}(1 - x_{mj}) = 1$. The only case satisfying this condition is $x_{ij} = 1$ and $x_{mj} = 1$. This constraint (7) only focuses on the $O\&D$ pair when the origin i and destination m are not the same. Constraint (8) limits the total flows on the secondary link. Building a capacitated hub j should satisfy two conditions: First, every hub j can choose at most one capacity level p in constraint (9); Second, the capacity of hub j should load the flows from both primary and secondary links. The left-hand side of constraint (10) represents the leftover capacity of hub j after the disruptions. The right-hand side of constraint (10) represents all the flows via hub j. Constraints (11) and (12) impose non-negativity on the integer decision variables.

The above model is formulated as a mixed-integer nonlinear programming (MINLP) due to the nonlinear terms $y_{imk}^{s}x_{ij}(1 - x_{mj})$, $y_{mik}^{s}x_{ij}(1 - x_{mj})$, $y_{imk}^{s}x_{ij}$ and $y_{mik}^{s}x_{ij}$ in objective (1) and the nonlinear terms $Z_{imk}x_{ij}(1 - x_{mj})$, $Z_{mik}x_{ij}(1 - x_{mj})$, $Z_{imk}x_{ij}$ and $Z_{mik}x_{ij}$ in constraints (7) and (10). Due to the computational complexity of MINLP, we propose a linear reformulation by using a linearization technique in which we introduce auxiliary nonnegative variables, additional variables, and big-M parameters.

To introduce the linearization technique, now let

$$u_{imj} = x_{ij}\left(1-x_{mj}\right) \quad i,\ j,\ m \in I.$$

$$O_{imjk}^{s} = y_{imk}^{s}x_{ij}\left(1-x_{mj}\right) \quad i,\ j,\ m \in I,\ k \in K \text{ and } s \in S.$$

$$G_{imjk} = Z_{imk}x_{ij}\left(1-x_{mj}\right) \quad i,\ j,\ m \in I \text{ and } k \in K.$$

$$N_{imjk}^{s} = y_{imk}^{s} x_{ij} \quad i,\ j,\ m \in I,\ k \in K \text{ and } s \in S.$$
$$E_{imjk} = Z_{imk} x_{ij} \quad i,\ j,\ m \in I \text{ and } k \in K. \tag{13}$$

Then we can replace the objective (1) with

$$\begin{aligned} \max \quad & \sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\sum_{s\in S} p_{imk}^{s} r_{imk} y_{imk}^{s} \\ & - \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}\sum_{s\in S} C_{1k}\left(p_{imk}^{s} f_{j0}\left\lceil \frac{O_{imjk}^{s}}{Q1}\right\rceil \right. \\ & \left. + p_{mik}^{s} f_{0j}\left\lceil \frac{O_{mijk}^{s}}{Q1}\right\rceil\right) \\ & - \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}\sum_{s\in S} C_{2k}\left(p_{imk}^{s} f_{ij}\left\lceil \frac{N_{imjk}^{s}}{Q2}\right\rceil \right. \\ & \left. + p_{mik}^{s} f_{ji}\left\lceil \frac{N_{mijk}^{s}}{Q2}\right\rceil\right) - \sum_{j\in I}\sum_{q\in W} F_{j}^{q} x_{jj} - F' \end{aligned} \tag{14}$$

and replace the nonlinear constraints (7) and (10) with

$$\begin{aligned} & \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil \frac{G_{imjk}}{Q1}\right\rceil + \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil \frac{G_{mijk}}{Q1}\right\rceil \\ & \quad \le No_{j0} x_{jj} \quad j \in I \end{aligned} \tag{15}$$

$$\begin{aligned} & \sum_{q\in W}\left(1-\alpha_{j}^{s}\right)\Gamma^{q} v_{j}^{q} \\ & \quad \ge \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil \frac{G_{imjk}}{Q1}\right\rceil + \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil \frac{G_{mijk}}{Q1}\right\rceil \\ & \quad + \sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\left\lceil \frac{E_{imjk}}{Q2}\right\rceil + \sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\left\lceil \frac{E_{mijk}}{Q2}\right\rceil \\ & \qquad j \in I,\ s \in S \end{aligned} \tag{16}$$

and also add the following additional constraints to make the linearization valid:

$$u_{imj} \ge x_{ij} - x_{mj} \quad i, j, m \in I \tag{17}$$

$$u_{imj} \ge 0 \quad i, j, m \in I \tag{18}$$

$$O_{imjk}^{s} \ge y_{imk}^{s} - M\left(1 - u_{imj}\right) \quad i, j, m \in I,\ k \in K,\ s \in S \tag{19}$$

$$O_{mijk}^{s} \ge y_{mik}^{s} - M\left(1 - u_{imj}\right) \quad i, j, m \in I,\ k \in K,\ s \in S \tag{20}$$

$$G_{imjk} \ge Z_{imk} - M\left(1 - u_{imj}\right) \quad i, j, m \in I,\ k \in K \tag{21}$$

$$G_{mijk} \ge Z_{mik} - M\left(1 - u_{imj}\right) \quad i, j, m \in I,\ k \in K \tag{22}$$

$$N_{imjk}^{s} \ge y_{imk}^{s} - M\left(1 - x_{ij}\right) \quad i, j, m \in I,\ k \in K,\ s \in S \tag{23}$$

$$N_{mijk}^{s} \ge y_{mik}^{s} - M\left(1 - x_{ij}\right) \quad i, j, m \in I,\ k \in K,\ s \in S \tag{24}$$

$$E_{imjk} \ge Z_{imk} - M\left(1 - x_{ij}\right) \quad i, j, m \in I,\ k \in K \tag{25}$$

$$E_{mijk} \ge Z_{mik} - M\left(1 - x_{ij}\right) \quad i, j, m \in I,\ k \in K \tag{26}$$

Here, this linear model as MILP formulation has objective (14) and constraints (2)-(6), (8)-(9), and (15)-(26).

4.2. Hybrid Two-Stage Stochastic Programming-Robust Optimization Model. To provide flexible decisions for this integrated problem, we combine the benefits of stochastic programming and robust optimization. Consequently, we provide a hybrid stochastic programming-robust optimization model based on the above nonlinear formulation in Section 4.1. The results can provide relative emphasis on the average-case and worst-case profits by introducing a pair of weights ω_1 and ω_2, where $\omega_1, \omega_2 \in [0, 1]$, $\omega_1 + \omega_2 = 1$.

This hybrid model is as follows:

$$\begin{aligned} \max \quad & \omega_1 \left\{ \sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\sum_{s\in S} p_{imk}^{s} r_{imk} y_{imk}^{s} - \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}\sum_{s\in S} C_{1k}\left(p_{imk}^{s} f_{j0}\left\lceil \frac{y_{imk}^{s}}{Q1}\right\rceil + p_{mik}^{s} f_{0j}\left\lceil \frac{y_{mik}^{s}}{Q1}\right\rceil\right)\left[x_{ij}\left(1 - x_{mj}\right)\right] \right. \\ & \left. - \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}\sum_{s\in S} C_{2k}\left(p_{imk}^{s} f_{ij}\left\lceil \frac{y_{imk}^{s}}{Q2}\right\rceil + p_{mik}^{s} f_{ji}\left\lceil \frac{y_{mik}^{s}}{Q2}\right\rceil\right) x_{ij} \right\} + \omega_2 \min_{s\in S}\left\{ \sum_{i\in I}\sum_{m\in I}\sum_{k\in K} r_{imk} y_{imk}^{s} \right. \\ & \left. - \sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K} C_{1k}\left(f_{j0}\left\lceil \frac{y_{imk}^{s}}{Q1}\right\rceil + f_{0j}\left\lceil \frac{y_{mik}^{s}}{Q1}\right\rceil\right)\left[x_{ij}\left(1 - x_{mj}\right)\right] \right. \end{aligned}$$

$$-\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}C_{2k}\left(f_{ij}\left\lceil\frac{y_{imk}^{s}}{Q2}\right\rceil+f_{ji}\left\lceil\frac{y_{mik}^{s}}{Q2}\right\rceil\right)x_{ij}\Bigg\}-\sum_{j\in I}\sum_{q\in W}F_{j}^{q}x_{jj}-F' \tag{27}$$

subject to (2)-(12).

We further simplify objective (27) by introducing an auxiliary variable Ψ ($\Psi \geq 0$). Then this model can be reformulated as follows:

$$\begin{aligned}\max\quad \omega_1\Bigg\{&\sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\sum_{s\in S}p_{imk}^{s}r_{imk}y_{imk}^{s}-\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}\sum_{s\in S}C_{1k}\left(p_{imk}^{s}f_{j0}\left\lceil\frac{y_{imk}^{s}}{Q1}\right\rceil+p_{mik}^{s}f_{0j}\left\lceil\frac{y_{mik}^{s}}{Q1}\right\rceil\right)\left[x_{ij}\left(1-x_{mj}\right)\right]\\&-\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}\sum_{s\in S}C_{2k}\left(p_{imk}^{s}f_{ij}\left\lceil\frac{y_{imk}^{s}}{Q2}\right\rceil+p_{mik}^{s}f_{ji}\left\lceil\frac{y_{mik}^{s}}{Q2}\right\rceil\right)x_{ij}\Bigg\}+\omega_2\Psi-\sum_{j\in I}\sum_{q\in W}F_{j}^{q}x_{jj}-F'\end{aligned} \tag{28}$$

subject to (2)-(12)

$$\begin{aligned}\Psi\leq&\sum_{i\in I}\sum_{m\in I}\sum_{k\in K}r_{imk}y_{imk}^{s}-\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}C_{1k}\left(f_{j0}\left\lceil\frac{y_{imk}^{s}}{Q1}\right\rceil+f_{0j}\left\lceil\frac{y_{mik}^{s}}{Q1}\right\rceil\right)\left[x_{ij}\left(1-x_{mj}\right)\right]\\&-\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}C_{2k}\left(f_{ij}\left\lceil\frac{y_{imk}^{s}}{Q2}\right\rceil+f_{ji}\left\lceil\frac{y_{mik}^{s}}{Q2}\right\rceil\right)x_{ij}\quad s\in S\end{aligned} \tag{29}$$

$$\Psi\geq 0 \tag{30}$$

5. Solution Approach

In this section, we introduce two approaches: (i) sample average approximation (SAA) and (ii) genetic algorithms. The details can be given in Sections 5.1 and 5.2.

5.1. Sample Average Approximation. SAA can be effective to deal with stochastic programming problems. The basic idea of SAA is to realize uncertain parameters by generating random samples and to approximate the expected value by the corresponding sample average function. The sample size can be much smaller than the number of scenarios in the true problem and SAA can converge exponentially fast to the true problem as the sample size increases [44]. SAA has been applied extensively in the network design ([45–47]).

In our paper, we apply the SAA scheme to solve the linear model in Section 4.1 by randomly generating the sample size L. Then the linear model in Section 4.1 is approximated by the following SAA problem:

$$\begin{aligned}\max\quad&\sum_{l\in L}\frac{1}{L}\Bigg\{\sum_{i\in I}\sum_{m\in I}\sum_{k\in K}r_{imk}y_{imk}^{l}\\&-\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}C_{1k}\left(f_{j0}\left\lceil\frac{O_{imjk}^{l}}{Q1}\right\rceil+f_{0j}\left\lceil\frac{O_{mijk}^{l}}{Q1}\right\rceil\right)\\&-\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{j\in I}\sum_{k\in K}C_{2k}\left(f_{ij}\left\lceil\frac{N_{imjk}^{l}}{Q2}\right\rceil+f_{ji}\left\lceil\frac{N_{mijk}^{l}}{Q2}\right\rceil\right)\Bigg\}\\&-\sum_{j\in I}\sum_{q\in W}F_{j}^{q}x_{jj}-F'\end{aligned} \tag{31}$$

subject to (2)-(4), (8)-(9), (15), (17)-(18), (21)-(22), (25)-(26), and

$$y_{imk}^{l}\leq d_{imk}^{l},\quad i,m\in I,\ k\in K,\ l\in L \tag{32}$$

$$y_{imk}^{l}\leq Z_{imk},\quad i,m\in I,\ k\in K,\ l\in L \tag{33}$$

$$\begin{aligned}\sum_{q\in W}\left(1-\alpha_{j}^{l}\right)\Gamma^{q}v_{j}^{q}\geq&\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil\frac{G_{imjk}}{Q1}\right\rceil\\&+\sum_{i\in I}\sum_{m\in I/\{i\}}\sum_{k\in K}\left\lceil\frac{G_{mijk}}{Q1}\right\rceil\\&+\sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\left\lceil\frac{E_{imjk}}{Q2}\right\rceil\\&+\sum_{i\in I}\sum_{m\in I}\sum_{k\in K}\left\lceil\frac{E_{mijk}}{Q2}\right\rceil\end{aligned} \tag{34}$$

$$j\in I,\ l\in L$$

$$O^l_{imjk} \geq y^l_{imk} - M\left(1 - u_{imj}\right) \quad i, j, m \in I,\ k \in K,\ l \in L \tag{35}$$

$$O^l_{mijk} \geq y^l_{mik} - M\left(1 - u_{imj}\right) \quad i, j, m \in I,\ k \in K,\ l \in L \tag{36}$$

$$N^l_{imjk} \geq y^l_{imk} - M\left(1 - x_{ij}\right) \quad i, j, m \in I,\ k \in K,\ l \in L \tag{37}$$

$$N^l_{mijk} \geq y^l_{mik} - M\left(1 - x_{ij}\right) \quad i, j, m \in I,\ k \in K,\ l \in L \tag{38}$$

$$x_{ij}, x_{jj}, v^q_j \in \{0, 1\}, \quad i, j, m \in I,\ q \in W,\ l \in L \tag{39}$$

$$\begin{aligned} &y^l_{imk}, y^l_{mik}, Z_{imk}, Z_{mik}, O^l_{imjk}, O^l_{mijk}, G_{imjk}, G_{mijk}, N^l_{imjk}, \\ &N^l_{mijk}, E_{imjk}, E_{mijk} \in Z_+, \quad i, m, j \in I,\ k \in K,\ l \in L \end{aligned} \tag{40}$$

$$u_{imj} \geq 0, \quad i, j, m \in I \tag{41}$$

5.2. Genetic Algorithms. As the size of the network increases, computational challenges appear when finding the optimal results. A lot of research has used GA to solve the capacitated hub location problem in reasonable computational time ([48–52]). The GA is inspired by natural biological evolution, which is introduced by [53]. In initial generation, the individuals are generated randomly. In addition, the individuals need to be selected according to their fitness in each population. The offspring can be obtained by using the crossover and mutation operators to deal with two parent individuals. This process can enhance the individuals better suited to the problem.

Here, we introduce a GA approach in [19]. First, this GA can enumerate all the possible individuals by the values of x_{ij} and x_{jj}. For the MINLP model in Section 4.1, the near optimal solutions can be obtained by CPLEX based on these individuals.

We briefly mention the procedure of this approach as follows:

Step 0 (initial population generation). Randomly generate a population of M chromosomes.

Step 0.1. A N-dimensional matrix is set from a uniform distribution with the interval $[0, 1]$, where N is the number of nodes. We name this matrix as matrix A.

Step 0.2. P highest entries from the diagonal of matrix A should be chosen. These P entries can be turned into 1 and the other unselected entries on this diagonal can be rewritten as 0. These entries equating to 1 on this diagonal mean that these nodes are chosen as the hubs, namely, $x_{jj} = 1$. $x_{jj} = 0$ indicates the entries equating to 0 on this diagonal.

Step 0.3. We observe the overlaps between all columns equating to 1 on this diagonal and all rows equating to 0 on the same diagonal. For these overlaps, we define the highest entry in each row as 1. On the contrary, the other entries in the same row of these overlaps should be 0. The entries equaling to 1 within these overlaps mean that these nodes are chosen as the non-hubs, namely, $x_{ij} = 1$. Otherwise $x_{ij} = 0$. The other entries in addition to the ones $x_{ij} = 1$, $x_{ij} = 0$, $x_{jj} = 1$, and $x_{jj} = 0$ can be rewritten as 0. Beyond that, if the overlaps in the same row have the same values, we choose the first highest one as $x_{ij} = 1$ from the left edge of this matrix.

Step 0.4. A 0-1 matrix can be obtained by these steps. We name this 0-1 matrix as matrix B.

Step 1 (fitness calculation). During reproduction, if the chromosome's ability can survive well in the present, it has a high probability of being selected for mating. The fitness value is used for measuring the survivability of the chromosome. Here, we set the reciprocal of the profit obtained from MINLP model in Section 4.1 as the fitness value of each individual.

Step 2 (a new population generation). Repeat the following three substeps until a new population is generated.

Step 2.1 (selection). The parents can be chosen from the population by random selection.

Step 2.2 (crossover). The crossover operator is used for producing the offsprings. The offspring can be generated by combining the parents with a crossover probability in following rules:

$$P_{new1} = bP_f + (1 - b)\, P_m \tag{42}$$

$$P_{new2} = (1 - b)\, P_f + bP_m \tag{43}$$

where P_{new1} and P_{new2} are the offsprings. P_f and P_m represent two different parents and each of them is represented by a N-dimensional matrix. b is a random N-dimensional matrix chosen uniformly over the interval $[0, 1]$. Each offspring is performed by a combination of two parents based on the above equalities (42) and (43) with the same matrix b.

Step 2.3 (mutation). The new offspring can be mutated at a mutation probability. The mutation operator can help to obtain a global flexible solution and to recover the good genetic codes which may lose from the above steps. We consider three mutation operators: Elimination, Transposition, and Conversion. In each iteration, only one of these mutation operators can be randomly selected. These three mutation operators are as follows:

(i) Elimination. The hub needs to be relocated in this operator. We name the highest entry on the diagonal of matrix A as place B. This place B needs to be subtracted from 1.

(ii) Transposition. This operator aims to change the links between the non-hubs and the hubs by transposing matrix A.

(iii) Conversion. A new network structure is established by this operator. A l-dimension matrix ($l < N$) is chosen from the bottom-right corner of matrix A, namely, matrix C. This random integer l is created from the multiplication of the nodes N and a random number drawn from a uniform distribution $[0, 0.5]$. We can generate a new l-dimension matrix of uniformly distributed between 0 and 1 to replace matrix C.

After using the mutation operator, new values of x_{jj} and x_{ij} are reattained from Step *0.2-0.4*. Accordingly, putting x_{jj} and x_{ij} into the MINLP model in Section 4.1 can obtain the values of other decision variables.

Step 3 (stopping criterion). If a stopping criterion is fulfilled, then terminate the iterations and output the best solution from the population. The stop conditions should be fulfilled by any one of these following two conditions:

(i) The best solution has not changed after 2 iterations.

(ii) The maximum running time of 8 hours has been reached.

6. Computational Experiments and Discussion

In this section, we conduct numerical experiments for our problem to verify the proposed mathematical models and to evaluate the performance of the used approaches. We use the SAA and the GA to solve the stochastic programming in Section 4.1. For the hybrid model in Section 4.2, we only apply the GA to obtain the results. First, we give all the parameters in Section 6.1. Then the results can be analyzed in Section 6.2.

6.1. Test Bed. For parameter setting, the nodes and the hubs (N, P) in this airline network are assumed to be 4 pairs: $(5, 2)$, $(10, 2)$, $(10, 3)$, and $(10, 5)$. We consider two cases for the capacity levels: (1)Case 1: every hub has two capacity levels to be chosen, 0.5×10^7 and 0.1×10^8. For the capacity level 0.5×10^7, the installation cost F_j^q of hub j is drawn from a uniform distribution $U(0.1 \times 10^7, 1.2 \times 10^6)$. For the capacity level 0.1×10^8, the installation cost F_j^q is drawn from a uniform distribution $U(1.2 \times 10^6, 1.4 \times 10^6)$. (2)Case 2: three capacity levels can be selected for each hub, 0.5×10^7, 0.1×10^8, and 1.5×10^7. The corresponding installation cost F_j^q is drawn from uniform distributions $U(0.1 \times 10^7, 1.2 \times 10^6)$, $U(1.2 \times 10^6, 1.4 \times 10^6)$, and $U(1.4 \times 10^6, 1.6 \times 10^6)$, respectively. The installation cost F' is drawn from a uniform distribution $U(1.2 \times 10^6, 2.4 \times 10^6)$.

Two fare classes ($K = 2$) involve a business class $k = 1$ and an economic class $k = 2$. For the primary link, the transportation cost C_{11} per unit distance per flight at the business class is drawn from $U(7, 14)$ and C_{12} at the economic class is γC_{11}, where $\gamma \sim U(0.5, 1)$. We consider the discount factor φ as 0.2 on the secondary link, which is $C_{21} = \varphi C_{11}$ and $C_{22} = \varphi C_{12}$. We assume that the distance f_{ij} from the non-hub to the hub is 10000 and from the hub to the central hub f_{j0} is 10000. The capacity No_{ij} on the primary link is drawn from a discrete uniform distribution $DU(1, 3)$. Meanwhile, the capacity No_{j0} on the secondary link is drawn from $DU(3, 6)$. The aircraft capacity Q_1 used on the primary link and Q_2 on the secondary link are 100 and 200, respectively. The ticket price at the economic class $k = 2$ is drawn from a continuous distribution $U(600, 3600)$. The price at the business class $k = 1$ is 1.5 higher than the one at $k = 2$. The demand d_{imk}^s for each fare class is drawn from the interval between low demands and high demands $[3.75, 6.25]$. 80% demand comes from the economic class $k = 2$ and 20% is from the business class $k = 1$.

About the scenarios, we set the number of scenarios S as 10. The probability of each scenario is the same, namely, $1/S$. In addition, the probability of crossover is 0.8 and the probability of mutation is 0.15. We consider the solution in 10 generations. The population size is 20. For the SAA, the sample size L is 50. Five-pair values of (ω_1, ω_2) are given: $(0, 1)$, $(0.2, 0.8)$, $(0.4, 0.6)$, $(0.6, 0.4)$, and $(0.8, 0.2)$. The weights $(1, 0)$ represent that the objective of our problem is to maximize the average profit. The weights $(0, 1)$ stand for maximizing the worst-case profit. The weight combinations $(0.2, 0.8)$, $(0.4, 0.6)$, $(0.6, 0.4)$, and $(0.8, 0.2)$ mean maximizing a weighted sum of average and worst case profits.

6.2. Computational Results and Analysis. In this section, we analysis the results of using the approaches in Section 5 on a series of instances. Computational experiments are implemented in a MacBook Pro using an Intel Core i7 3GHz CPU with 8GB RAM, running macOS High Sierra 10.13.3 operating system. Furthermore, all experiments are coded in Matlab R2017b which itself calls YALMIP R20171121 to run through CPLEX 12.8.0. For all the instances, one run is conducted for each approach, according to the settings in Section 6.1. Then all the results are analyzed as follows.

First, we compare the results of the SAA and the GA for the stochastic programming problem in Table 1. We consider the instances with two capacity levels $W = 2$ and with three capacity levels $W = 3$. In Table 1, the results of $W = 2$ are showed from rows 4 to 7. The results of $W = 3$ are reported between rows 9 and 12. The node N and the hub P are listed in the column (N, P). The columns under SAA represent the profit and the computation time of this algorithm. It is the same for the columns under GA which represent the profit and the computation time. The unit of computation time is 1 s. In terms of profit indicator, the SAA achieves slightly larger values than the GA in 6 among 8 instances. It is reasonable because the value of the best individual within a population can be regarded as the solution of our problem for the GA. However, the SAA considers the mean value of all the samples. Even though the value of one of the samples is the worst, the profit can't sharply curtailed. In particular, the worst sample has less influence on the mean when more samples are considered. In terms of running time indicator, the SAA performs better than the GA in 5 out of 8 instances. These results indicate that the SAA is superior

Table 1: Computational results obtained when solving the stochastic programming model.

(N, P)	SAA		GA	
	Profit	Time	Profit	Time
$W = 2$				
(5,2)	6210033.41	1801.90	6174857.72	7103.15
(10,2)	6667668.11	15148.28	6609379.61	13287.96
(10,3)	8933666.56	15791.17	8868768.82	>28800
(10,5)	13575857.79	23226.39	13557851.53	31107.05
$W = 3$				
(5,2)	9497307.55	483.19	9452119.92	1105.79
(10,2)	8807660.40	13565.30	8755399.7	13387.39
(10,3)	12622118.92	3534.22	12573458.11	14513.20
(10,5)	20263852.37	37609.61	20307462.00	16223.70

Table 2: Computational results obtained when solving the hybrid model.

(N, P)	(ω_1, ω_2)				
	$(0, 1)$	$(0.2, 0.8)$	$(0.4, 0.6)$	$(0.6, 0.4)$	$(0.8, 0.2)$
(5,2)	6181317.90	6180025.86	6178733.83	6177441.79	6176149.75
(10,2)	6618414.92	6616607.86	6614800.80	6612993.74	6611186.67
(10,3)	8876628.71	8875056.73	8873484.75	8871912.78	8870340.80
(10,5)	13563862.65	13562660.42	13561458.20	13560255.98	13559053.75

to the GA in computation time. This table well illustrates the SAA performs better than the GA in most instances.

Second, we apply the GA to the hybrid problem in Section 4.2 for the instances with $W = 2$ in Table 2. All the computations are done in 8 hours. In Table 2, the column (N, P) represents the number of the nodes N and the hubs P. The columns under the weights (ω_1, ω_2) represent the profit considering five pairs of weights. For every column, the profit gets larger as the scale of the network becomes larger. These results reflect that the profit is relevant to the instance size. In particular, larger profit can be obtained when this network with the same number of the nodes has more hubs. It is reasonable because the hub can facilitate the transport and create more business opportunities. As such, more profits can be obtained in the hub network. For each row, the profit is decreasing as ω_1 increases and ω_2 decreases. The maximum is obtained at $(0, 1)$ and the minimum is at $(0.8, 0.2)$. These results show that ω_2 has a greater influence on the profit than ω_1. More specifically, the worst-case value plays an essential role in the profit compared to the average-case value. Hence, this analysis shows that the decision maker can make flexible decisions by putting relative emphasis on either worst-case or average-case value in the profit according to the practical issues.

In addition, we also collect the worst-case values by applying the GA for the hybrid model in Table 3. Single run has been carried out for the instances with two capacity levels. The worst-case values of Ψ are shown under five pairs of weights. For each column, Ψ manifests a strong increase as the instance size increases. For the instances with the same number of the nodes and different number of the hubs, Ψ becomes larger in the network with more hubs. For every row, the values of Ψ are the same even with different weights. It is reasonable because Ψ is relevant to the network size, the scenario, and the capacity level in our problem. If the values of the above-mentioned parameters are fixed, then the value of Ψ can't be changed under any case of the weights.

For illustrative purposes, we plot all the results from Tables 1 and 2 in Figure 2. This figure comprises four subfigures representing the profits in four instances: (1) $N = 5$ and $P = 2$, (2) $N = 10$ and $P = 2$, (3) $N = 10$ and $P = 3$, and (4) $N = 10$ and $P = 5$. The marker "■" represents the solutions computed by SAA for the stochastic programming model in Table 1. The marker "●" illustrates the profits obtained from the GA for the stochastic programming model in Table 1. The marker "▲" indicates the results computed by the GA for the hybrid model in Table 2. In each subfigure, the x-axis denotes the approach and the weights. The y-axis indicates the profit. We observe that the maximum profits can be obtained by using the SAA for the stochastic programming model and the minimum profits are collected by the GA for the stochastic programming model in each subfigure. The values for the hybrid model are between the values computed by SAA and by the GA for the stochastic programming model. The values of SAA are greater than the ones with the weights $(0, 1)$. In fact, the values with the weights $(0, 1)$ are very restrictive because the results are the worst-case values from a robust optimization model. As such, the values of the SAA are more restrictive in solving the proposed instances. Different from the case under the weights $(0, 1)$, the values at $(0.8, 0.2)$ are from solving a hybrid model which considers less impact from the worst-case value so that the results are less restrictive. We also conclude that the values obtained by the GA for the stochastic programming model are less than the values under the weight $(0.8, 0.2)$. Hence, the hybrid model outperforms the stochastic programming model in these instances.

Table 3: Computational results for Ψ.

(N, P)	Ψ				
	(0, 1)	(0.2, 0.8)	(0.4, 0.6)	(0.6, 0.4)	(0.8, 0.2)
(5,2)	30023.80	30023.80	30023.80	30023.80	30023.80
(10,2)	50786.85	50786.85	50786.85	50786.85	50786.85
(10,3)	60270.74	60270.74	60270.74	60270.74	60270.74
(10,5)	66517.83	66517.83	66517.83	66517.83	66517.83

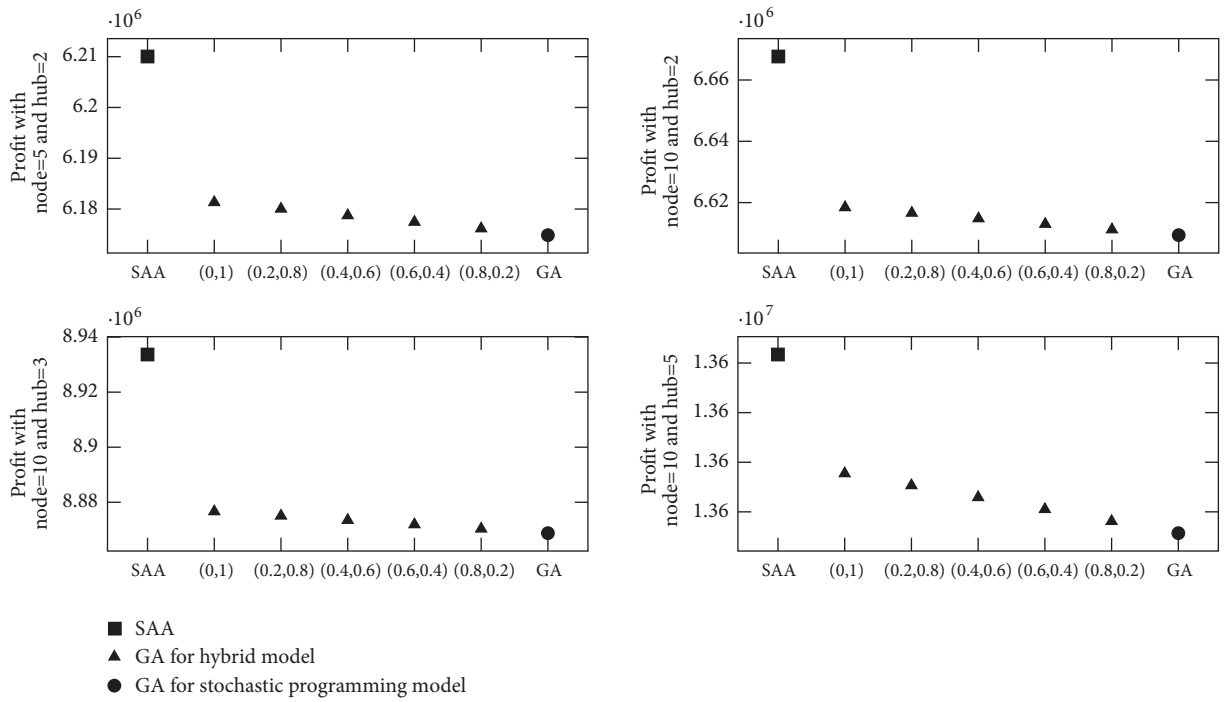


Figure 2: Computational solutions obtained when solving two models.

7. Conclusions

This paper addressed an integration of single allocation capacitated p-hub location and revenue management in air transportation. The hubs with multiple capacity levels and the disruption were considered. We modeled this problem as a two-stage stochastic programming formulation and a two-stage hybrid stochastic programming-robust optimization model to maximize the profit in a star-star network. In the first stage, the location of the hub, the link between the hub and the non-hub, the capacity level of the hub, and the protection level of tickets for different booking classes were determined. The booking limit of the tickets could be obtained in the second stage. We used a set of discrete scenarios to represent uncertain demand and network disruption at the hub. The SAA and the GA were applied to solve the stochastic programming model. For the hybrid model, we also used the GA. In the numerical experiments, we obtained the computational results when solving the stochastic programming model and the hybrid model. The performance of the SAA and the GA for the stochastic programming model was compared. Obviously, the SAA outperformed the GA in computing time and high-quality results. Accordingly, we analyzed the factors influencing the results for the hybrid model. The network size and the weight played a role in the profit and the results of Ψ. Furthermore, we demonstrated that the hybrid model outperforms the stochastic programming model.

Future studies could be enriched from these perspectives: (1) there should be an extension to a competitive hub location problem by considering the factor of market competition, and (2) an exact algorithm could be developed to reduce the computation time.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant number 71371140].

References

[1] A. De Juniac, IATA Annual Review, 2018.

[2] "Airport capacity: striking the right deal," https://www.airlines.iata.org/analysis/airport-capacity-striking-the-right-deal.

[3] G. Laporte, S. Nickel, and F. S. da Gama, "Introduction to location science," in *Location Science*, pp. 1–18, Springer, Cham, Switzerland, 2015.

[4] https://www.iata.org/training/Pages/revenue-management.aspx.

[5] K. Pak and N. Piersma, "Overview of OR techniques for airline revenue management," *Statistica Neerlandica*, vol. 56, no. 4, pp. 480–496, 2002.

[6] I. Correia, S. Nickel, and F. Saldanha-da-Gama, "Single-assignment hub location problems with multiple capacity levels," *Transportation Research Part B: Methodological*, vol. 44, no. 8-9, pp. 1047–1066, 2010.

[7] F. Fotuhi and N. Huynh, "Reliable intermodal freight network expansion with demand uncertainties and network disruptions," *Networks and Spatial Economics*, vol. 17, no. 2, pp. 405–433, 2017.

[8] Z. Liang, F. Xiao, X. Qian et al., "A column generation-based heuristic for aircraft recovery problem with airport capacity constraints and maintenance flexibility," *Transportation Research Part B: Methodological*, vol. 113, pp. 70–90, 2018.

[9] V. Gabrel, C. Murat, and A. Thiele, "Recent advances in robust optimization: an overview," *European Journal of Operational Research*, vol. 235, no. 3, pp. 471–483, 2014.

[10] J. Dalal and H. Üster, "Combining worst case and average case considerations in an integrated emergency response network design problem," *Transportation Science*, vol. 52, no. 1, pp. 171–188, 2018.

[11] M. E. O'Kelly, "The location of interacting hub facilities," *Transportation Science*, vol. 20, no. 2, pp. 92–106, 1986.

[12] A. Alibeyg, I. Contreras, and E. Fernández, "Hub network design problems with profits," *Transportation Research Part E: Logistics and Transportation Review*, vol. 96, pp. 40–59, 2016.

[13] R. Neamatian Monemi, S. Gelareh, S. Hanafi, and N. Maculan, "A co-opetitive framework for the hub location problems in transportation networks," *Optimization*, vol. 66, no. 12, pp. 2089–2106, 2017.

[14] A. Alibeyg, I. Contreras, and E. Fernindez, "Exact solution of hub network design problems with profits," *European Journal of Operational Research*, vol. 266, no. 1, pp. 57–71, 2018.

[15] G. Taherkhani and S. A. Alumur, "Profit maximizing hub location problems," *Omega*, 2018.

[16] A. Lüer-Villagra and V. Marianov, "A competitive hub location and pricing problem," *European Journal of Operational Research*, vol. 231, no. 3, pp. 734–744, 2013.

[17] C.-C. Lin and S.-C. Lee, "Hub network design problem with profit optimization for time-definite LTL freight transportation," *Transportation Research Part E: Logistics and Transportation Review*, vol. 114, pp. 104–120, 2018.

[18] H. Tikani, M. Honarvar, and Y. Z. Mehrjerdi, "Joint optimization of star P-hub median problem and seat inventory control decisions considering a hybrid routing transportation system," *International Journal of Supply and Operations Management*, vol. 3, no. 3, pp. 1–25, 2016.

[19] H. Tikani, M. Honarvar, and Y. Zare Mehrjerdi, "Developing an integrated hub location and revenue management model considering multi-classes of customers in the airline industry," *Computational and Applied Mathematics*, vol. 37, no. 3, pp. 3334–3364, 2018.

[20] W.-C. Chiang, J. C. H. Chen, and X. Xu, "An overview of research on revenue management: current issues and future research," *International Journal of Revenue Management*, vol. 1, no. 1, pp. 97–128, 2007.

[21] K. Littlewood, "Forecasting and control of passenger bookings," Airline Group International Federation of Operational Research Societies Proceedings, vol. 12, pp. 95-117, 1972.

[22] P. P. Belobaba, "Survey paper—airline yield management an overview of seat inventory control," *Transportation Science*, vol. 21, no. 2, pp. 63–73, 1987.

[23] F. Glover, R. Glover, J. Lorenzo, and C. McMillan, "The passenger-mix problem in the scheduled airlines," *Interfaces*, vol. 12, no. 3, pp. 73–80, 1982.

[24] K. Talluri and G. Van Ryzin, "An analysis of bid-price controls for network revenue management," *Management Science*, vol. 44, no. 11, pp. 1577–1593, 1998.

[25] P.-S. You, "Dynamic pricing in airline seat management for flights with multiple flight legs," *Transportation Science*, vol. 33, no. 2, pp. 192–206, 1999.

[26] S. Kunnumkal and K. Talluri, "On a piecewise-linear approximation for network revenue management," *Mathematics of Operations Research*, vol. 41, no. 1, pp. 72–91, 2016.

[27] W. Grauberger and A. Kimms, "Revenue management under horizontal and vertical competition within airline alliances," *Omega* , vol. 59, pp. 228–237, 2016.

[28] W. Grauberger and A. Kimms, "Airline revenue management games with simultaneous price and quantity competition," *Computers & Operations Research*, vol. 75, pp. 64–75, 2016.

[29] H.-Z. He, "Structural properties in a hub-to-hub network revenue management problem," *Journal of the Operations Research Society of China*, vol. 4, no. 4, pp. 503–516, 2016.

[30] J. I. McGill and G. J. van Ryzin, "Revenue management: research overview and prospects," *Transportation Science*, vol. 33, no. 2, pp. 233–256, 1999.

[31] K. T. Talluri and G. J. van Ryzin, *The Theory And Practice of Revenue Management*, vol. 68, Springer Science and Business Media, 2006.

[32] J. F. Campbell and M. E. O'Kelly, "Twenty-five years of hub location research," *Transportation Science*, vol. 46, no. 2, pp. 153–169, 2012.

[33] R. Z. Farahani, M. Hekmatfar, A. B. Arabani, and E. Nikbakhsh, "Hub location problems: a review of models, classification, solution techniques, and applications," *Computers & Industrial Engineering*, vol. 64, no. 4, pp. 1096–1109, 2013.

[34] A. T. Ernst and M. Krishnamoorthy, "Solution algorithms for the capacitated single allocation hub location problem," *Annals of Operations Research*, vol. 86, pp. 141–159, 1999.

[35] I. Correia, S. Nickel, and F. Saldanha-da-Gama, "The capacitated single-allocation hub location problem revisited: a note on a classical formulation," *European Journal of Operational Research*, vol. 207, no. 1, pp. 92–96, 2010.

[36] J. F. Campbell, "Integer programming formulations of discrete hub location problems," *European Journal of Operational Research*, vol. 72, no. 2, pp. 387–405, 1994.

[37] D. Skorin-Kapov, J. Skorin-Kapov, and M. O'Kelly, "Tight linear programming relaxations of uncapacitated p-hub median problems," *European Journal of Operational Research*, vol. 94, no. 3, pp. 582–593, 1996.

[38] I. Correia, S. Nickel, and F. Saldanha-da-Gama, "Hub and spoke network design with single-assignment, capacity decisions and balancing requirements," *Applied Mathematical Modelling*, vol. 35, no. 10, pp. 4841–4851, 2011.

[39] A. M. Hörhammer, "Dynamic hub location problems with single allocation and multiple capacity levels," in *Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS '14)*, pp. 994–1003, IEEE, January 2014.

[40] S. A. Alumur, S. Nickel, F. Saldanha-da-Gama, and Y. Seerdin, "Multi-period hub network design problems with modular capacities," *Annals of Operations Research*, vol. 246, no. 1-2, pp. 289–312, 2016.

[41] S. A. Alumur, S. Nickel, B. Rohrbeck, and F. Saldanha-da-Gama, "Modeling congestion and service time in hub location problems," *Applied Mathematical Modelling*, vol. 55, pp. 13–32, 2018.

[42] C.-C. Lin and S.-H. Chen, "An integral constrained generalized hub-and-spoke network design problem," *Transportation Research Part E*, vol. 44, no. 6, pp. 986–1003, 2008.

[43] R. J. Wets, "Stochastic programs with fixed recourse: the equivalent deterministic program," *SIAM Review*, vol. 16, pp. 309–339, 1974.

[44] S. Ahmed, A. Shapiro, and E. Shapiro, "The sample average approximation method for stochastic programs with integer recourse," Submitted for publication, pp. 1-24, 2002.

[45] T. Santoso, S. Ahmed, M. Goetschalckx, and A. Shapiro, "A stochastic programming approach for supply chain network design under uncertainty," *European Journal of Operational Research*, vol. 167, no. 1, pp. 96–115, 2005.

[46] G. R. Patil and S. V. Ukkusuri, "Sample average approximation technique for flexible network design problem," *Journal of Computing in Civil Engineering*, vol. 25, no. 3, pp. 254–262, 2011.

[47] B. Ayvaz, B. Bolat, and N. Aydin, "Stochastic reverse logistics network design for waste of electrical and electronic equipment," *Resources, Conservation & Recycling*, vol. 104, pp. 391–404, 2015.

[48] Z. Stanimirovic, "Solving the capacitated single allocation hub location problem using genetic algorithm," in *Recent Advances in Stochastic Modeling and Data Analysis*, pp. 464–471, 2007.

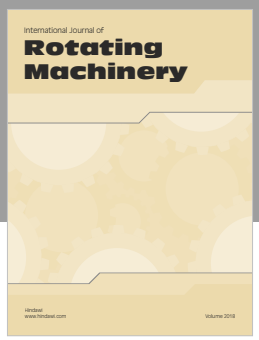
[49] Z. Stanimirović, "A genetic algorithm approach for the capacitated single allocation p-hub median problem," *Computing and Informatics*, vol. 29, no. 1, pp. 117–132, 2012.

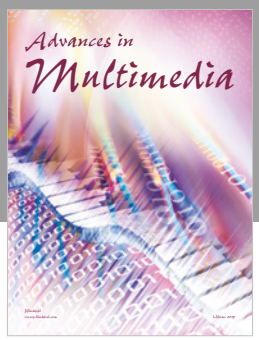
[50] C.-C. Lin, J.-Y. Lin, and Y.-C. Chen, "The capacitated *p*-hub median problem with integral constraints: an application to a Chinese air cargo network," *Applied Mathematical Modelling: Simulation and Computation for Engineering and Environmental Systems*, vol. 36, no. 6, pp. 2777–2787, 2012.

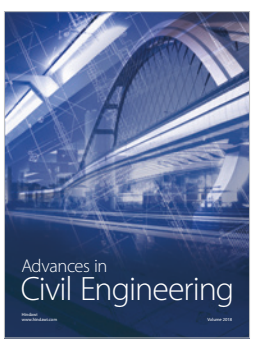
[51] M. Bashiri, M. Mirzaei, and M. Randall, "Modeling fuzzy capacitated p-hub center problem and a genetic algorithm solution," *Applied Mathematical Modelling*, vol. 37, no. 5, pp. 3513–3525, 2013.

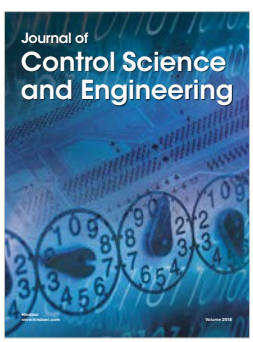
[52] J. Zheng, Q. Meng, and Z. Sun, "Liner hub-and-spoke shipping network design," *Transportation Research Part E: Logistics and Transportation Review*, vol. 75, pp. 32–48, 2015.

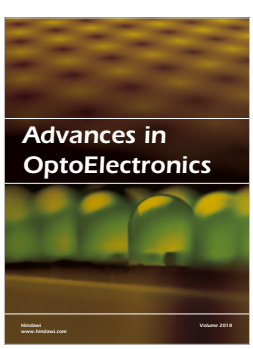
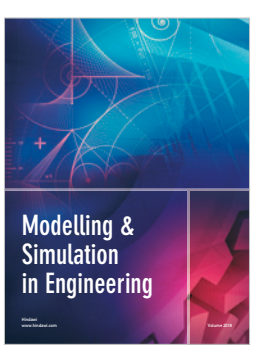
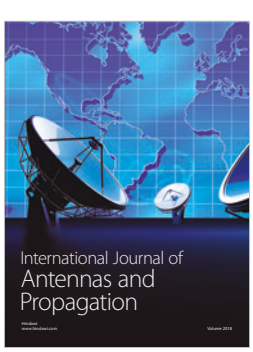
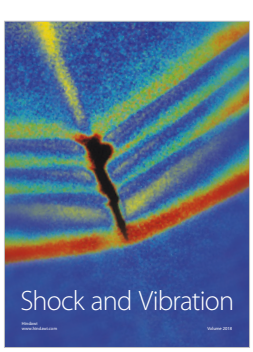
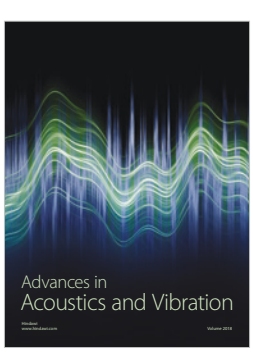
[53] J. H. Holland, "Genetic algorithms," *Scientific American*, vol. 267, no. 1, pp. 66–72, 1992.









Hindawi

Submit your manuscripts at
www.hindawi.com