IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 28, 2020, accepted March 23, 2020, date of publication April 3, 2020, date of current version April 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2985377

Robust Hub Location Problem With Flow-Based Set-Up Cost

SHUXIA LI, CHENHAO FANG, AND YIFAN WU
School of Business, East China University of Science and Technology, Shanghai 200237, China
Corresponding author: Yifan Wu (yifanwu@ecust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 71471062 and Grant 71431004, in part by the Natural Science Foundation of Shanghai under Grant 18ZR1409400, and in part by the Ministry of Education Humanities and Social Science Research Planning Fund Project under Grant 18YJAZH046.

ABSTRACT Hub location problems are network design problems on the level of strategic decision-making processes. During strategic planning, input data, such as flows and set-up costs, are not always known in advance. Hence, decisions have to be made in an uncertain environment. In this paper, two sources of uncertainty are considered: the flows from origins to destinations and the set-up costs of hubs. A robust optimization formulation is proposed for both single and multiple allocation cases, in which the flow between each pair of nodes is assumed to be uncertain and correlated. In addition, the set-up cost of a hub is related to the total flow through the hub. Nonlinear integer program models are presented for both single and multiple allocation cases, and they are solved using CPLEX. Computational tests using the Civil Aeronautics Board and Australian Post datasets are provided. The numerical results suggest that the robust optimization strategy locates more hubs than in the deterministic case with a relatively small cost increase, and the total cost of the robust solution calculated for the multiple allocation case is marginally lower than that for the single allocation case. The robust optimization strategy is proven to be effective for protecting the solution against the worst case for different uncertain parameters.

INDEX TERMS Hub location problem, nonlinear integer program, robust optimization.

I. INTRODUCTION

Hub location problems have attracted the attention of many scholars for more than 20 years due to their wide range of applications, particularly in air transportation and telecommunications. Hub location problems involve locating isolated or connected hub facilities, allocating non-hub nodes to hubs, and routing from origins to destinations. A transportation cost exists between each pair of nodes, and this cost is based on time or distance. Because the transportation cost between each pair of hubs is discounted, presumably to make use of the economies of scale of consolidated transportation, a significant reduction can be obtained by routing the traffic flow via hubs. Therefore, compared to classical modes of transportation, a hub network can fulfill the transportation between a large number of origins and destinations using fewer links.

Most studies of hub location problems are based on deterministic circumstances. In practice, however, hub location problems are network design problems on the level of strategic decision-making process, which may depend on uncertain parameters. For example, in postal and cargo services, the flows between each pair of nodes vary greatly with uncertain changes in demand. Furthermore, although the set-up costs may be estimated in advance, this estimate may not be accurate due to factors such as the price of the materials, the cost of land resources and the labor cost, which are affected by the total flows through the hubs. Results based on the deterministic model may lead to dramatic cost growth for certain cases. This paper presents a robust hub location optimization model with a flow-based set-up cost for both single and multiple allocation cases under uncertain flow. The uncertain flow is assumed to be correlated and affected by few sources of uncertainty. Each source of uncertainty can be represented by a random variable that belongs to an ellipsoidal uncertainty set. In addition, the set-up cost of a hub is related to the total flow through the hub. In this way, the robust optimization strategy is designed to protect the solution against the worst case for different uncertain scenarios. For both single and multiple

The associate editor coordinating the review of this manuscript and approving it for publication was Yanbo Chen.

allocation cases, nonlinear integer programs are provided, and they can be solved using the optimization software CPLEX.

A. RELATED WORK

Since the seminal paper by O'Kelly [1], [2], a considerable amount of work has been conducted in this area. Problems with different features have been addressed, including the p-hub location problem, p-hub median location problem, p-hub center location problem and p-hub covering location problem. According to Farahani *et al.* [3], p-hub median location problems are multiple allocation p-hub location problems. P-hub center location problems [4]–[7] are defined based on the minimax criterion, in which the maximum cost of the origin-destination pair is minimized. In p-hub covering location problems [8]–[11], hub facilities are considered to be located such that the origin-destination pair of two non-hub nodes is covered by a pair of hub nodes. In addition, the allocation mode of non-hub nodes to hubs is an important aspect of hub location problems. Two types of allocation modes are often considered: single allocation and multiple allocation. In the former case, each non-hub node can be allocated to only one hub. Studies such as [12]–[15] are based on this mode. In the latter case, each non-hub node can be allocated to multiple hubs. Ernst and Krishnamoorthy [16], Ebery *et al.* [17], Kratica [18], Bolanda *et al.* [19], and Mayer and Wagner [20] have performed a large amount of work in this field.

Optimization under uncertainty is always one of the hot topics in decision making problem. Many works deal with uncertainty in supply chain management [21]–[24]. In many hub location decision problems, it is important to take the uncertainty of factors, such as flows and set-up costs, into consideration. Regarding flow, the information will become obsolete as time elapses, leading to random fluctuations and seasonal changes in flows [25]–[29]. Some works such as Alumur *et al.* [30] consider mixed uncertainties of flow and set-up cost. In addition, uncertain time is another aspect of uncertainty [31], [32]. Few papers explore the relationship between set-up cost and the total flow through a hub. Different from the case in Alumur *et al.* [30], in this study, set-up cost will be uncertain and unpredictable due to the uncertainty of the flow.

To address uncertainty, two main optimization approaches are proposed. One is stochastic optimization, and the other is robust optimization. Stochastic optimization relies on the distribution of the uncertain parameters [33], [34], whereas robust optimization is a distribution-free approach that aims to find the worst-case scenario with respect to a predefined uncertainty set. The interest in robust optimization was revived in the 1990s. Ben-Tal *et al.* [35], Ben-Tal and Nemirovski [36], Ben-Tal [37] introduced a number of formulations and provided mathematical analyses of linear and convex optimizations. Merakli and Yaman [38] adopted both the hose model and the hybrid model to study the robust uncapacitated multiple allocation p-hub median problem under polyhedral demand uncertainty. Talbi and Todosijevic [39] introduced a new way to quantify the robustness of a solution in the presence of flow uncertainties, and this method can cope with any realization of the number of changes that may occur. Miskovic and Stanimirovic [40] introduced a robust variant of the uncapacitated multiple allocation p-hub center problem by considering flow variations with unknown distributions. Abbasi-Parizi *et al.* [41] studied a minimax regret hub location problem in a fuzzy-stochastic environment incorporating risk factors such as availability, security, delay time, environmental guidelines and regional air pollution. Others such as Bertsimas *et al.* [42], Bertsimas and Sim [43], Chen *et al.* [44], Zetina *et al.* [45] and Ghaffari-Nasab *et al.* [46], Ghaffari-Nasab [47] have also conducted a vast amount of work in this direction.

B. CONTRIBUTIONS

In this paper, the flow between each pair of nodes, which is assumed to be uncertain and correlated, can be expressed as a linear combination with a known mean and several independent random variables [27]. Each independent random variable represents a source of uncertainty, leading to flow disturbance from the mean flow. In addition, the set-up cost of a hub is assumed to be proportional to the total flow through the hub. A robust optimization strategy is proposed for both single and multiple allocation cases. To the best of the authors' knowledge, Reference [43] is the first work addressing correlated data, where data uncertainty affects the violation of constraints. However, in our robust formulations, flow uncertainty is incorporated into the objective function, which is suitable for problems without uncertain parameters in the constraints. Furthermore, the robust optimization models for both single and multiple allocation cases are presented to protect the solution against the worst case for different uncertain scenarios.

The remainder of this paper is organized as follows. Deterministic formulations are first presented in Section 2. Section 3 provides the robust formulations for both single allocation cases and multiple allocation cases. Computational tests are presented in Section 4. Finally, the conclusions are provided in the last section.

II. DETERMINISTIC FORMULATIONS

The robust formulations of the hub location problem are developed based on the deterministic models provided by Skorin-Kapov *et al.* [48] for multiple allocation and by Ernst and Krishnamoorthy [12] for single allocation. The deterministic single allocation and multiple allocation models are shown in the following section.

A. NOTATIONS

f_{ij} is the flow between each pair of nodes;

c_{ij} is the transportation cost per unit of flow;

α is the discount factor for making use of the economies of scale from consolidated transportation.

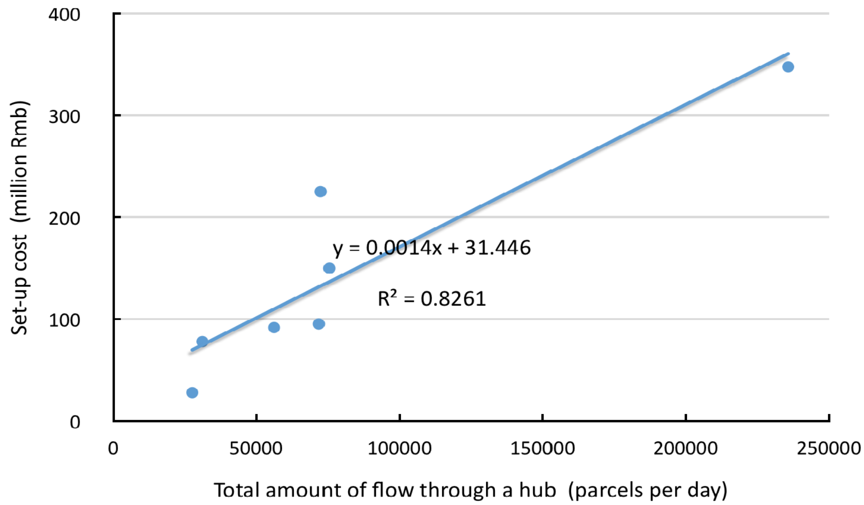


FIGURE 1. Linear regression analysis.

B. SINGLE ALLOCATION HUB LOCATION PROBLEM WITH FLOW-BASED SET-UP COST

In single allocation cases, every non-hub node can be allocated to only one hub. Here, the variable z_{ik} indicates whether non-hub node i is allocated to hub k. The variable x_{ijkl} represents the routing strategy, i.e., whether the flow to be transported from node i to node j passes through hubs k and l. The set-up cost is assumed to be proportional to the total flow through a hub. According to the data of a Chinese logistics company, linear regression analysis is conducted to investigate the relationship between these factors, as shown in Fig. 1. The value of R^2 is 0.826, which indicates a good linear correlation. O_k represents the total flow through hub k; thus, $O_k = \sum_i \sum_j z_{ik} f_{ij}$. The mathematical programming formulation of the single allocation hub location problem is given below.

$$\min \sum_i \sum_j \sum_k \sum_l (c_{ik} + \alpha c_{kl} + c_{lj}) x_{ijkl} f_{ij} + \sum_k z_{kk} (a \sum_i \sum_j z_{ik} f_{ij} + b) \tag{1}$$

$$s.t. \sum_k z_{ik} = 1, \quad \forall i \tag{2}$$

$$z_{ik} \leqslant z_{kk}, \quad \forall i, k \tag{3}$$

$$\sum_k x_{ijkl} = z_{jl}, \quad \forall i, j, l \tag{4}$$

$$\sum_l x_{ijkl} = z_{ik}, \quad \forall i, j, k \tag{5}$$

$$x_{ijkl}, z_{ik} \in \{0, 1\} \tag{6}$$

Here, a and b are the parameters of the linear function. Objective function (1) is the minimization of the total cost, including the transportation cost and set-up cost. Constraint (2) ensures that each non-hub node can be allocated to only one hub. Constraint (3) ensures that a non-hub node can be allocated to a hub if and only if this hub is already established. Constraints (4) and (5) guarantee that a path from the origin to the destination through a pair of hubs exists if and only if these two non-hub nodes are allocated to these two hubs separately. Constraint (6) is the binary decision variable constraint.

C. MULTIPLE ALLOCATION HUB LOCATION PROBLEM WITH FLOW-BASED SET-UP COST

In multiple allocation, every non-hub node can be allocated to more than one hub. Here, variable y_k represents the location of the hub facility, while x_{ijkl} is the same as that in the single allocation case. Similarly, the total flow through hub k can be written as $O_k = \sum_i \sum_j \sum_l x_{ijkl} f_{ij}$. The mathematical programming formulation of the multiple allocation hub location problem is given below.

$$\min \sum_i \sum_j \sum_k \sum_l (c_{ik} + \alpha c_{kl} + c_{lj}) x_{ijkl} f_{ij} + \sum_k y_k (a \sum_i \sum_j \sum_l x_{ijkl} f_{ij} + b) \tag{7}$$

$$s.t. \sum_k \sum_l x_{ijkl} = 1, \quad \forall i, j \tag{8}$$

$$\sum_l x_{ijkl} \leqslant y_k, \quad \forall i, j, k \tag{9}$$

$$\sum_k x_{ijkl} \leqslant y_l, \quad \forall i, j, l \tag{10}$$

$$x_{ijkl}, y_k \in \{0, 1\} \tag{11}$$

Similar to the single allocation case, objective function (7) is the minimization of the total cost. Constraint (8) ensures that each path from the origin to the destination passes through a pair of hubs. Constraints (9) and (10) ensure that each path from the origin to the destination through a pair of hubs exists if and only if these two hubs are already established. Constraint (11) is the binary decision variable constraint.

III. ROBUST FORMULATIONS

According to Shahabi and Unnikrishnan [27], the flow is considered to be uncertain and correlated. In particular, it is affected by few sources of uncertainty. The uncertain flow is composed of a known mean and several independent random variables, as shown in (12).

$$\tilde{f}_{ij} = \bar{f}_{ij} + \sum_{m=1}^{M} b_{ijm}\tilde{u}_m, \quad \forall i, j \in N \tag{12}$$

Here, $\bar{f}_{ij}$ represents the mean flow to be transported from node i to node j. b_{ijm} is the m^{th} weight corresponding to random variable $\tilde{u}_m$. $\tilde{u}_m$ refers to the primitive uncertainty variable, which is independent and symmetrically distributed according to the following three assumptions:

(i) $E(\tilde{u}_m) = 0, \forall m$;

(ii) $\|\tilde{u}\|_\infty = \max\{|\tilde{u}_1|, |\tilde{u}_2|, \ldots, |\tilde{u}_m|\} \leqslant 1$;

(iii) $\tilde{u}_m$ are all independent $\forall m$.

In other words, the uncertain flow can be expressed by an affine function of the nominal mean value and m independent uncertainty sources. The sources of uncertainty may include changes in economy, policy and population, competition, seasonal fluctuations and other uncertain factors.

Under affne data perturbation, the worst case uncertainty set is a parallelotope, and the robust optimization strategy considering the worst case for all sources of uncertainty will lead to an over-conservative solution. To avoid this case, the random variable is assumed to belong to an ellipsoidal set $U_\Omega = \{\tilde{u}_m, \|\tilde{u}\|_2 \leqslant \Omega\}$, where Ω is called the uncertainty budget [44]. The parameter Ω varies the size of the uncertainty set radially from the central point, such that $U_\Omega \subseteq U_{\Omega'} \subseteq W$ for all $\Omega_{max} \geq \Omega' \geq \Omega \geq 0$. Here, the worst case uncertainty set W is the convex support of the uncertain data, and Ω_{max} is the worst case budget of uncertainty. For the robust counterpart problem, the probability of feasibility guarantee is defined as follows:

$$p = 1 - exp(-\Omega^2/2) \tag{13}$$

In fact, p reveals the relationship between the possibility of uncertainty protection and the size of the uncertainty set. A larger uncertainty budget ensures a more conservative solution against the uncertainties, whereas a smaller uncertainty budget will lead to a less conservative solution. The choice of $\Omega = 0$ is an unprotected solution against uncertainties, while an increase in the uncertainty budget will increase the conservatism of the solution. Note that even the maximum uncertainty budget is substantially smaller than the worst-case budget, which is equal to $\sqrt{m}$.

By defining $c_{ijkl} = c_{ik} + \alpha c_{kl} + c_{lj}$, the robust formulation for the single allocation hub location problem with a flow-based set-up cost can be formulated as a min-max problem as follows:

$$\min_{x,z} \max_{\|\tilde{u}\|_2 \leqslant \Omega} \sum_i \sum_j \sum_k \sum_l c_{ijkl}x_{ijkl}(\bar{f}_{ij} + \sum_m b_{ijm}\tilde{u}_m) + \sum_k z_{kk}[a \sum_i \sum_j z_{ik}(\bar{f}_{ij} + \sum_m b_{ijm}\tilde{u}_m) + b] \tag{14}$$

Here, Ω is the uncertainty budget that controls the level of conservatism of the robust solution to achieve a desired level of robustness.

Because the inner maximization problem is based on the random variable $\tilde{u}_m$, the objective function can be rewritten as (15):

$$\min_{x,z}\{\sum_i \sum_j \sum_k \sum_l c_{ijkl}x_{ijkl}\bar{f}_{ij} + \sum_k z_{kk}[a \sum_i \sum_j z_{ik}\bar{f}_{ij} + b] + \max_{\|\tilde{u}\|_2 \leqslant \Omega}\{\sum_i \sum_j \sum_k \sum_l \sum_m c_{ijkl}x_{ijkl}b_{ijm}\tilde{u}_m + a \sum_i \sum_j \sum_k \sum_m z_{kk}z_{ik}b_{ijm}\tilde{u}_m\}\} \tag{15}$$

This formulation is a nonlinear integer program, which is difficult to solve. To transform this model into a computational model, the inner maximization problem needs to be solved first.

$$S = \max_{\|\tilde{u}\|_2 \leqslant \Omega} \sum_i \sum_j \sum_k \sum_l \sum_m c_{ijkl}x_{ijkl}b_{ijm}\tilde{u}_m + a \sum_i \sum_j \sum_k \sum_m z_{kk}z_{ik}b_{ijm}\tilde{u}_m \tag{16}$$

To solve this maximization problem, a Lagrangian relaxation approach is used, as presented in [27]. Given the optimal value of $\tilde{u}_m$, the robust formulation can be rewritten as a conic quadratic program. Let δ be the Lagrangian multiplier with $\|\tilde{u}\|_2 \leqslant \Omega$; then, the Lagrangian function can be rewritten as

$$L(\tilde{u}_m, \delta) = \sum_i \sum_j \sum_k \sum_l \sum_m c_{ijkl}x_{ijkl}b_{ijm}\tilde{u}_m + a \sum_i \sum_j \sum_k \sum_m z_{kk}z_{ik}b_{ijm}\tilde{u}_m - \delta(\Omega - \|\tilde{u}\|_2) \tag{17}$$

The first-order condition of (17) is

$$\frac{\partial L}{\partial \tilde{u}_m} = \sum_i \sum_j \sum_k \sum_l c_{ijkl}x_{ijkl}b_{ijm} + a \sum_i \sum_j \sum_k z_{kk}z_{ik}b_{ijm} - \delta\frac{\tilde{u}_m}{\|\tilde{u}\|_2} = 0 \tag{18}$$

$$\frac{\partial L}{\partial \delta} = \Omega - \|\tilde{u}\|_2 = 0 \tag{19}$$

From the first-order condition, if δ is considered to be positive, then the following equations can be obtained:

$$\Omega = \|\tilde{u}\|_2 \tag{20}$$

$$\tilde{u}_m = \frac{\|\tilde{u}\|_2}{\delta}(\sum_i \sum_j \sum_k \sum_l c_{ijkl}x_{ijkl}b_{ijm} + a \sum_i \sum_j \sum_k z_{kk}z_{ik}b_{ijm}) \tag{21}$$

$$\|\tilde{u}\|_2 = \frac{\Omega}{\delta}(\sum_m(\sum_i\sum_j\sum_k\sum_l c_{ijkl}x_{ijkl}b_{ijm} + a\sum_i\sum_j\sum_k z_{kk}z_{ik}b_{ijm})^2)^{1/2} \quad (22)$$

$$\delta = (\sum_m(\sum_i\sum_j\sum_k\sum_l c_{ijkl}x_{ijkl}b_{ijm} + a\sum_i\sum_j\sum_k z_{kk}z_{ik}b_{ijm})^2)^{1/2} \quad (23)$$

Integrating all the above equations, the optimal value of $\tilde{u}_m$ can be written as

$$\tilde{u}_m = \Omega(\sum_i\sum_j\sum_k\sum_l c_{ijkl}x_{ijkl}b_{ijm} + a\sum_i\sum_j\sum_k z_{kk}z_{ik}b_{ijm}) \,/\, (\sum_m(\sum_i\sum_j\sum_k\sum_l c_{ijkl}x_{ijkl}b_{ijm} + a\sum_i\sum_j\sum_k z_{kk}z_{ik}b_{ijm})^2)^{1/2} \quad (24)$$

Thus, the inner maximization problem can be written as

$$S = \Omega(\sum_m(\sum_i\sum_j\sum_k\sum_l c_{ijkl}x_{ijkl}b_{ijm} + a\sum_i\sum_j\sum_k z_{kk}z_{ik}b_{ijm})^2)^{1/2} \quad (25)$$

TABLE 1. Initial value of perturbation b_{ijm}.

	Low	Medium	High
b_{ijm}	$0.3f_{ij}$	$0.6f_{ij}$	$0.9f_{ij}$

Incorporating the solution of the maximization problem into the minimization problem, the robust single allocation hub location problem with a flow-based set-up cost can be converted into the following minimization problem:

$$\min_{x,z} \sum_i\sum_j\sum_k\sum_l c_{ijkl}x_{ijkl}\bar{f}_{ij} + \sum_k z_{kk}[a\sum_i\sum_j z_{ik}\bar{f}_{ij} + b] + \Omega(\sum_m(\sum_i\sum_j\sum_k\sum_l c_{ijkl}x_{ijkl}b_{ijm} + a\sum_i\sum_j\sum_k z_{kk}z_{ik}b_{ijm})^2)^{1/2}$$
$$s.t.(2)-(6) \quad (26)$$

Similar to single allocation problem, the robust multiple allocation hub location problem with a flow-based set-up cost can be written as

$$\min_{x,y} \sum_i\sum_j\sum_k\sum_l c_{ijkl}x_{ijkl}\bar{f}_{ij} + \sum_k y_k(a\sum_i\sum_j\sum_l \bar{f}_{ij}x_{ijkl} + b)$$

TABLE 2. Single allocation results with $\Omega = 1.5$.

α	Robust hub locations	Robust cost	Base hub locations	Base cost
Low uncertainty				
0.2	12,20	2283.78	12,20	1979.83
0.4	12,20	2407.17	12,20	2080.55
0.6	12,20	2529.15	5	2155.03
0.8	5	2569.17	5	2155.03
Medium uncertainty				
0.2	4,12,17	2563.54	12,20	1979.83
0.4	12,20	2733.78	12,20	2080.55
0.6	12,20	2878.17	5	2155.03
0.8	5	2983.29	5	2155.03
High uncertainty				
0.2	4,12,17	2814.94	12,20	1979.83
0.4	4,12,18	3039.98	12,20	2080.55
0.6	12,20	3227.19	5	2155.03
0.8	12,20	3382.76	5	2155.03

TABLE 3. Multiple allocation results with $\Omega = 1.5$.

α	Robust hub locations	Robust cost	Base hub locations	Base cost
Low uncertainty				
0.2	12,20	2277.8	12,20	1974.94
0.4	12,20	2371.47	12,20	2051.41
0.6	12,20	2450.59	12,20	2116
0.8	12,20	2503.19	5	2155.04
Medium uncertainty				
0.2	12,17,21	2542.6	12,20	1974.94
0.4	12,20	2691.53	12,20	2051.41
0.6	12,20	2785.19	12,20	2116
0.8	12,20	2847.45	5	2155.04
High uncertainty				
0.2	12,17,21	2790.75	12,20	1974.94
0.4	4,12,17	2969.52	12,20	2051.41
0.6	4,12,17	3119.59	12,20	2116
0.8	12,20	3191.7	5	2155.04

$$+\Omega(\sum_m(\sum_i\sum_j\sum_k\sum_l c_{ijkl}x_{ijkl}b_{ijm} + a\sum_i\sum_j\sum_k\sum_l x_{ijkl}y_k b_{ijm})^2)^{1/2}$$
$$s.t.(8)-(11) \tag{27}$$

IV. COMPUTATIONAL TESTS

In this section, computational tests with the Civil Aeronautics Board (CAB25) and Australian Post (AP) datasets are conducted to evaluate the solution of the robust formulation for single and multiple allocation problems. The quality of the solution obtained by the robust optimization approach in terms of hub locations and robust cost is studied first. Then, the sensitivities of the parameters of the robust model are analyzed with respect to the discount factor, uncertainty budget and uncertainty level. All models are coded in the AMPL platform and solved using CPLEX 12.6 with an Intel Core i5 processor running at 2.7 GHz and 8 GB of RAM under a 64-bit OS X Yosemite operating system.

A. INITIALIZATION OF INPUT PARAMETERS

The CAB and AP datasets presented in the OR-Library are commonly used in the hub location literature. The CAB data contain costs and flows based on the airline passenger interactions among 25 US cities, and the AP dataset contains the information about 200 zip districts in Australia, together with a computer code for reducing the size of the set by grouping districts and the flows between them. Similar to [30], each flow is divided by the total flow such that the total flow is equal to 1.

For the linear function of the set-up cost and the total flow through a hub, a and b are set as constants. For the CAB dataset, $a = 350$ and $b = 314.46$.

For uncertainty set construction, the uncertain flow is assumed to be dependent on four independent random variables with weight parameter b_{ijm}. To reduce the computational time, $b_{ijm} = b_m f_{ij}$, where b_m is a constant in [0,1].

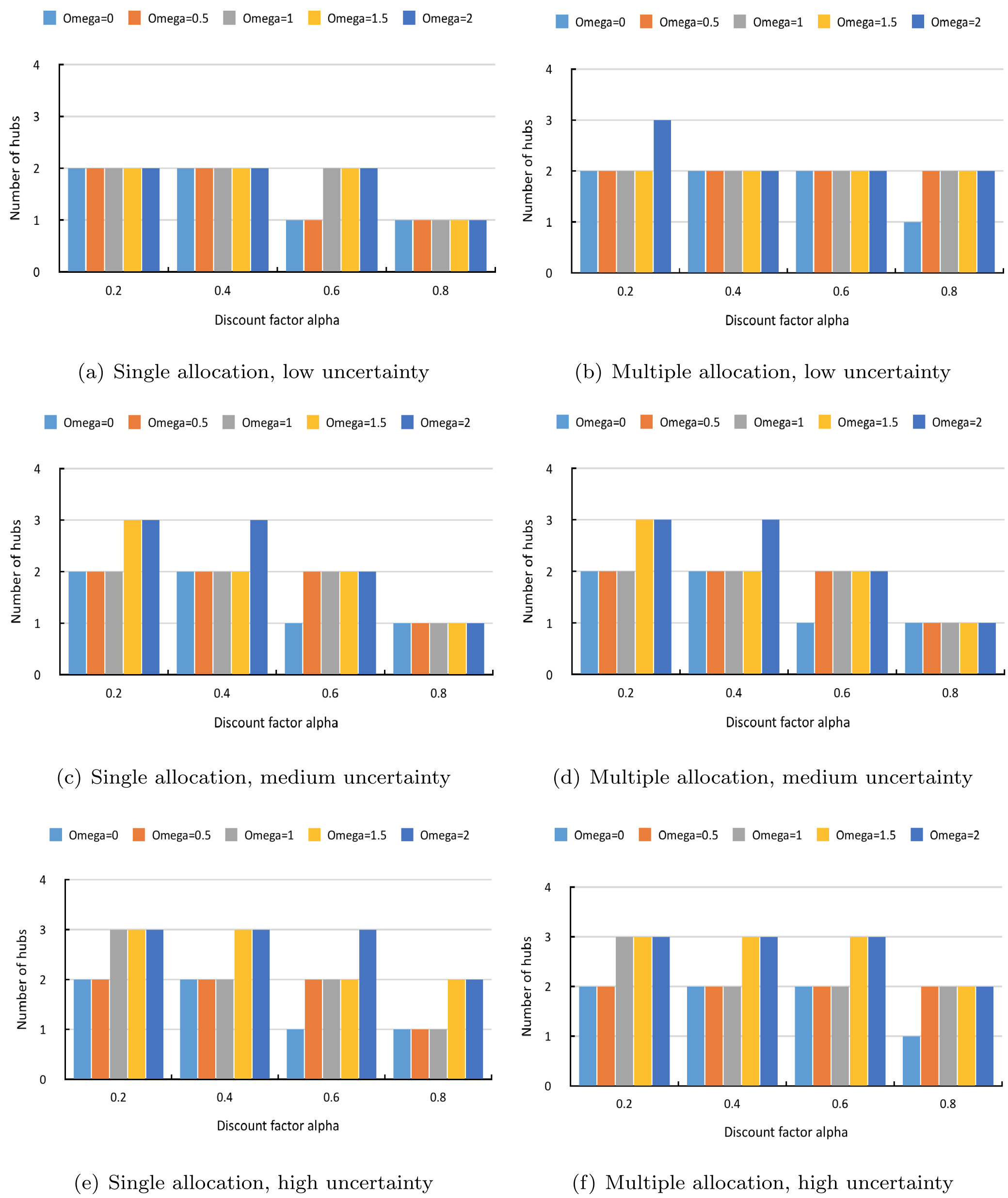
Three levels of flow uncertainty, namely, low, medium and high, are tested in this paper. Table 1 shows the initial values of b_{ijm} for the three levels of flow uncertainty.

B. NUMERICAL RESULTS

The first part of the numerical results aims to analyze the location decisions under different discount factors and uncertainty levels. In this part, the CAB dataset is employed, and $\alpha \in \{0.2, 0.4, 0.6, 0.8\}$. The uncertainty budget Ω is fixed at 1.5. The deterministic solutions are achieved at mean flow and reported as the base case. The robust hub locations, robust cost, base hub locations and base cost for the single and multiple allocation cases are shown in Tables 2 and 3, respectively.

The computational times of all cases in this part are less than 150 seconds. According to these two tables, three basic insights are obtained:

(1) Compared with the deterministic case, the number of hubs and the total cost tend to be greater in the robust case.



(a) Single allocation, low uncertainty

(b) Multiple allocation, low uncertainty

(c) Single allocation, medium uncertainty

(d) Multiple allocation, medium uncertainty

(e) Single allocation, high uncertainty

(f) Multiple allocation, high uncertainty

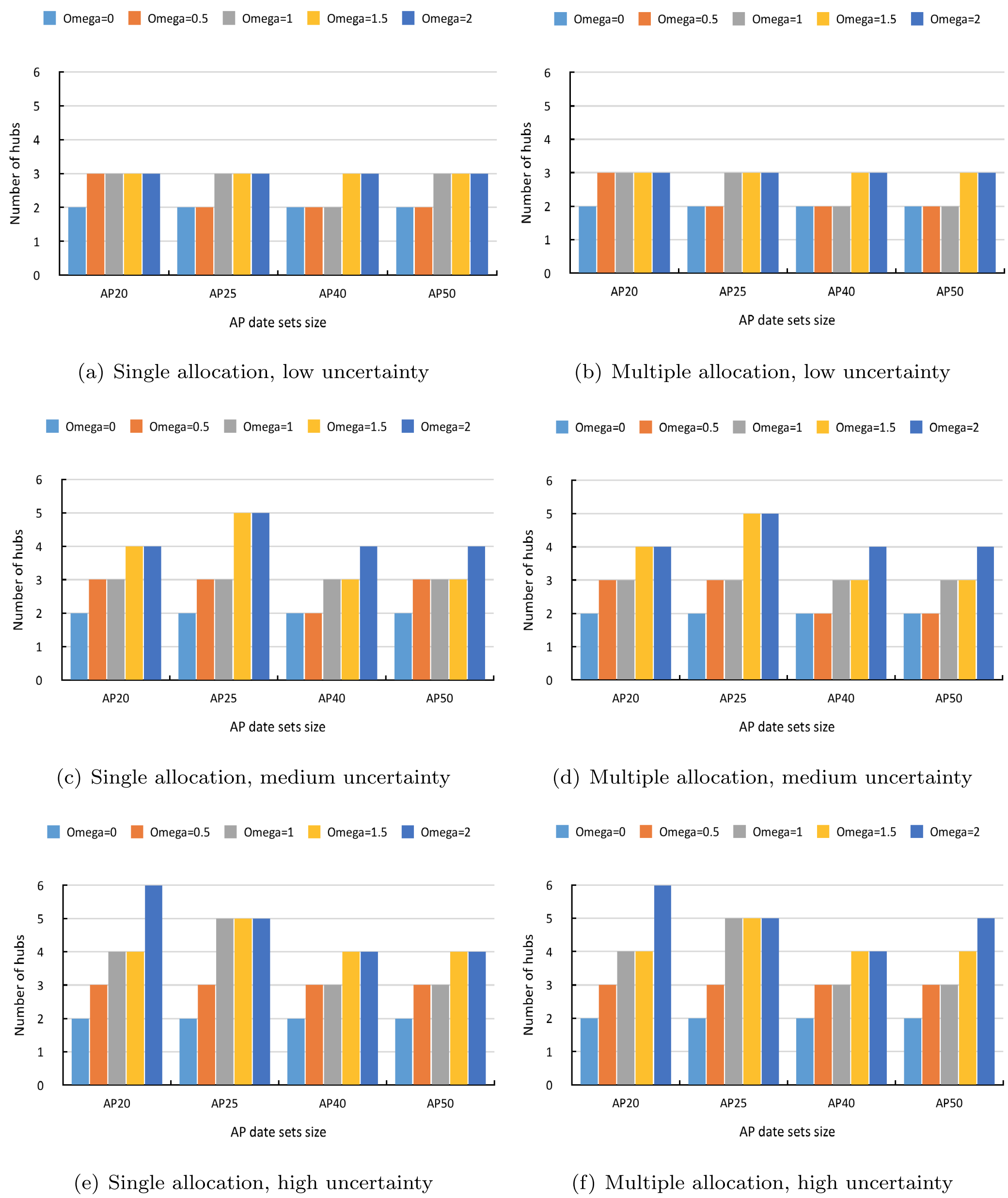
FIGURE 2. Number of robust hub locations for both single allocation and multiple allocation with different uncertainty levels (CAB dataset).

(2) With increasing uncertainty level, the number of hubs increases.

(3) When the discount factor increases, fewer hubs tend to be selected.

The first insight can be explained by the fact that the robust optimization strategy tends to protect the solution against the worst cases of different uncertain parameters. In the worst case, flows will be exaggerated due to the uncertainty budget, and the total cost will increase as a result. Consequently, the robust optimization strategy attempts to locate more hubs than the deterministic case. Similarly, when the uncertainty level increases, the robust optimization strategy locates more hubs to respond to the exaggerated flows. For example, in the single allocation case, the robust approach locates two hubs for the low uncertainty level, two hubs for the medium uncertainty level and three hubs for the high uncertainty level under the condition of $\alpha = 0.4$. This result also indicates that the robust approach locates the same hubs as the base case but with different costs under some circumstances. For instance, node 12 and node 20 are selected as hubs for both the robust case and the base case in some scenarios. This special phenomenon is caused by the uncertainty in the flow-based set-up cost.

In addition, the locations of hubs for the single and multiple allocation cases are not always the same. In single allocation problems, each non-hub node can be allocated to only one hub. However, in multiple allocation problems, each non-hub node can be allocated to two or more hubs. Thus, the location



(a) Single allocation, low uncertainty

(b) Multiple allocation, low uncertainty

(c) Single allocation, medium uncertainty

(d) Multiple allocation, medium uncertainty

(e) Single allocation, high uncertainty

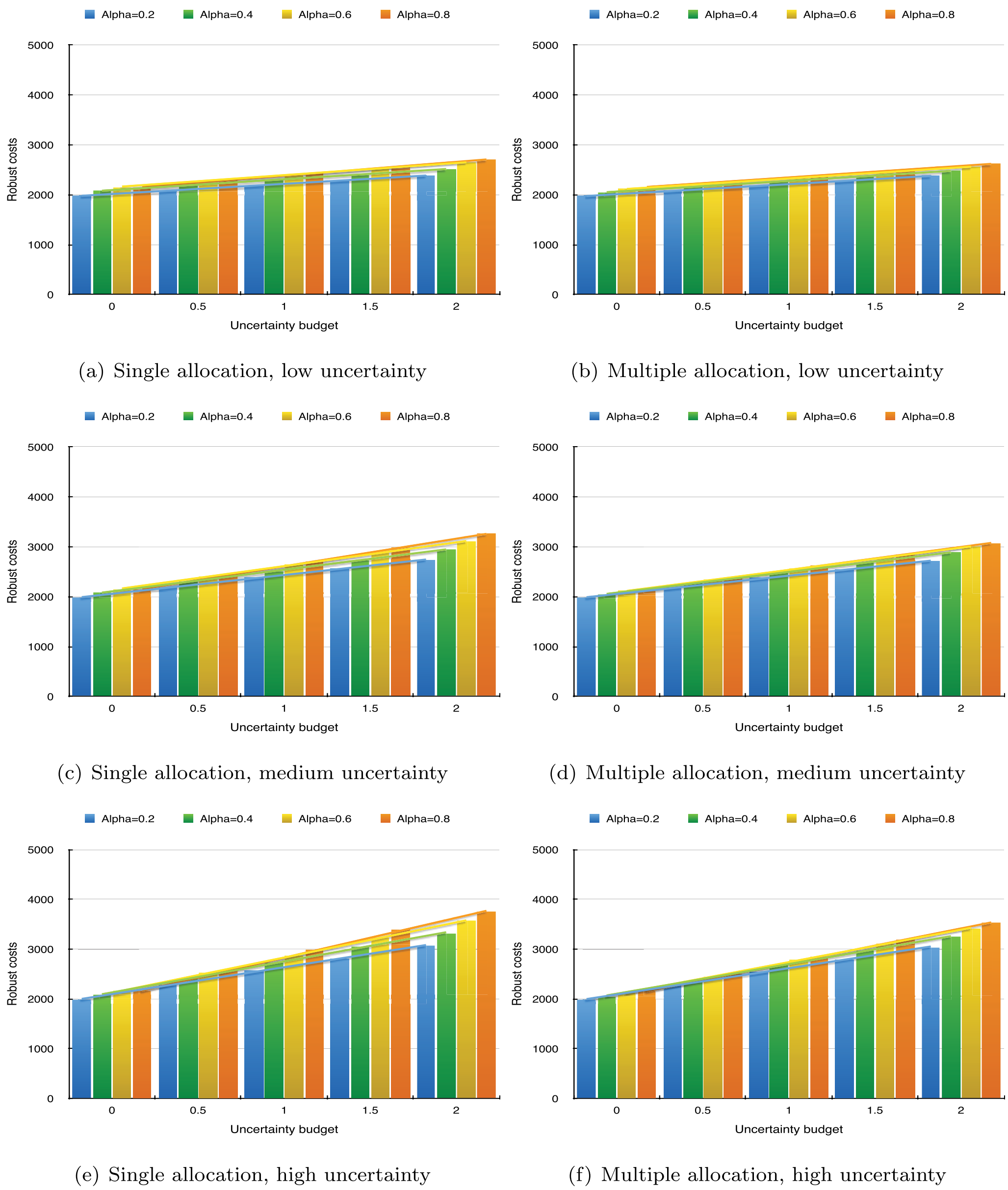
(f) Multiple allocation, high uncertainty

FIGURE 3. Number of robust hub locations for both single allocation and multiple allocation with different uncertainty levels (AP dataset, $\alpha = 0.6$).

decisions for single and multiple allocation problems may be different. More importantly, the cost of multiple allocation is less than that of single allocation.

Another trend in all the test cases is that the number of selected hubs tends to decrease with the increase in the discount factor. For example, 3 and 2 hubs are selected for multiple allocation in the robust case with the medium uncertainty level for $\alpha = 0.2$ and 0.4, respectively. In addition, in some cases, the number of robust hub locations is 1, which means that no hub arc exists in this scenario. For example, when the discount factor is 0.8, only one hub will be selected for single allocation in cases with low and medium uncertainty levels. In these cases, the hub does not benefit from the economies of scale of consolidated transportation. These special solutions demonstrate that, with the increase in the discount factor, in some circumstances, the decrease in the transportation cost of locating one more hub cannot compensate for the increase in the flow-based set-up cost of locating one more hub.

The goal of the next set of numerical experiments is to illustrate the effect of the uncertainty budget on hub location decisions. To this end, CAB25 and AP datasets with four different sizes of problem or number of nodes ($|N| = 20$, 25, 40 and 50) are employed. For the AP test problems, the computation time rises to around 550 seconds with the increase in the size of the problem. The single allocation and multiple allocation models are solved with uncertainty budgets varying from 0 to 2 with a step size of 0.5, which means that the protection probability varies from 0 to 86.5%.



(a) Single allocation, low uncertainty
(b) Multiple allocation, low uncertainty
(c) Single allocation, medium uncertainty
(d) Multiple allocation, medium uncertainty
(e) Single allocation, high uncertainty
(f) Multiple allocation, high uncertainty

FIGURE 4. Robust cost for both single allocation and multiple allocation with different uncertainty levels (CAB dataset).

The numbers of robust hub locations for both single and multiple allocation cases with different uncertainty levels are shown in Fig. 2 and Fig. 3, respectively. The result reveals that the number of hub locations tends to increase with increasing uncertainty budget in all the test scenarios. A higher uncertainty budget leads to more conservative solutions that locate more hubs to achieve a higher level of robustness.

Furthermore, the relative variations in the number of selected hubs for single allocation with different uncertainty budgets when compared to multiple allocation is studied. For the CAB25 test problem, it is found that the numbers of selected hubs for single and multiple allocation problems are similar if the factor of α is eliminated. Discarding the solutions under $\alpha = 0.6$ and 0.8 when economies of scale are not obvious, most of the results subject to the same uncertainty budget for both allocation modes locate the same number of hubs regardless of the realized uncertainty in demand. When $\Omega = 2$, an additional hub is selected for multiple allocation compared to single allocation for $\alpha = 0.2$ with a low uncertainty level. To take the demographical distribution and demand pattern into consideration, experiments based on AP datasets are also conducted. In this case, α is set to 0.6. As shown in Fig. 3, there are few cases in which the number of selected hubs for multiple allocation is different from that for single allocation, although the locations of the hubs may differ. Thus, the uncertainty budget hedges against uncertain demand regardless of the allocation mode.

Then, the costs of robust and deterministic hub location decisions for different uncertainty levels are compared based on both the CAB and AP datasets. The uncertainty budget varies from 0 to 2 with a step size of 0.5, and a discount factor α is taken from the set $\Lambda = \{0.2, 0.4, 0.6, 0.8\}$. Fig. 4 presents the costs of the robust hub location decision to uncertainty budgets for the single and multiple allocation cases with different uncertainty levels on the CAB dataset. Note that $\Omega = 0$ for the base case, where the cost is evaluated at mean flow. However, in other cases, the costs are evaluated with different uncertainty levels.

As shown in Fig. 4, the robust cost increases with the uncertainty budget. The solution tends to be more conservative to accommodate the deviation from the mean flow, resulting in a higher cost location decision. From the trend lines of each figure, it is clear that, with the increase in the discount factor, the robust cost increases more rapidly. Taking the result of the high uncertainty level for single allocation as an example, the gradients of the lines are 270.42, 311.49, 353.23 and 400.89. This result is expected because the unit cost from hub to hub increases with the increase in discount factor, leading to the increase in transportation cost. In addition, a higher discount factor means fewer hubs, resulting in more flow through each hub, which contributes to the rapid increase in the flow-based set-up cost. Therefore, special attention needs to be paid for estimating the discount factor if the uncertainty level is high. Furthermore, the maximum increase in the robust cost for the two allocation modes is 26% compared to the 30% increase in flow at the low uncertainty level, 51% compared to 60% at the medium uncertainty level and 74% compared to 90% at the high uncertainty level. It is clear that the increase in robust cost is more sensitive at the high level of uncertainty than at the other two levels.

According to the above analysis, it is evident that, in all cases, the robust cost calculated for the single allocation case is slightly higher than that for the multiple allocation case. Consequently, multiple allocation is a better choice in terms of total cost savings. Moreover, the robust optimization strategy is an effective measure for protecting the solution against the worst case of different uncertain parameters, particularly at the high uncertainty level.

V. CONCLUSION

In this paper, a robust optimization strategy is proposed for single and multiple allocation hub location problems under uncertain flows. The goal of the robust optimization model is to protect the solution against the worst cases of different uncertain parameters, which belong to an ellipsoidal uncertainty set. Nonlinear integer program models for the single and multiple allocation cases are presented, considering a flow-based set-up cost. Computational tests based on the CAB dataset are provided. The numerical results indicate that the robust strategy locates more hubs with higher total costs than the deterministic case. Moreover, a higher level of uncertainty and a smaller discount factor lead to the location of more hubs. The results also show that the robust cost and the number of hubs increase as the level of protection increases. Compared to that of multiple allocation, the robust cost calculated for the single allocation case is slightly higher, which indicates that multiple allocation is potentially a better strategy. Finally, the results prove that the robust optimization strategy is an effective measure for hedging the solution against uncertainty in demand, particularly at the high uncertainty level.

In contrast to Shahabi and Unnikrishnan [27], in this paper, the set-up cost of a hub is considered to be uncertain and to be affected by the total flow. Fewer hubs are needed than the number reported in the literature due to the flow-based set-up cost. However, the robust hub location decision does not merely depend on the allocation modes and set-up costs of the hubs. The capacity of a hub and the transportation time will have a significant effect on the hub location decision. Consequently, further study is warranted to obtain a more comprehensive understanding. In addition, the robust strategy finds only solutions for a predefined number of changes that are not necessarily optimal for other uncertain scenarios in consideration of the computational tractability. Therefore, it is worth considering new robust strategies that provide one or more solutions that are optimal for any flow or optimal for some uncertainties while not worse for others.

REFERENCES

[1] M. E. O'Kelly, "Activity levels at hub facilities in interacting networks," *Geograph. Anal.*, vol. 18, no. 4, pp. 343–356, 1986.
[2] M. E. O'Kelly, "A quadratic integer problem for the location of interacting hub facilities," *Eur. J. Oper. Res.*, vol. 32, pp. 393–404, 1987.
[3] R. Z. Farahani, M. Hekmatfar, A. B. Arabani, and E. Nikbakhsh, "Hub location problems: A review of models, classification, solution techniques, and applications," *Comput. Ind. Eng.*, vol. 64, no. 4, pp. 1096–1109, Apr. 2013.
[4] E. Hult, H. Jiang, and D. Ralph, "Exact computational approaches to a stochastic uncapacitated single allocation p-hub center problem," *Comput. Optim. Appl.*, vol. 59, nos. 1–2, pp. 185–200, Oct. 2014.
[5] K. Yang, Y. Liu, and G. Yang, "An improved hybrid particle swarm optimization algorithm for fuzzy p-hub center problem," *Comput. Ind. Eng.*, vol. 64, no. 1, pp. 133–142, Jan. 2013.
[6] A. M. Campbell, T. J. Lowe, and L. Zhang, "The p-hub center allocation problem," *Eur. J. Oper. Res.*, vol. 176, no. 2, pp. 819–835, Jan. 2007.
[7] A. T. Ernst, H. Hamacher, H. Jiang, M. Krishnamoorthy, and G. Woeginger, "Uncapacitated single and multiple allocation p-hub center problems," *Comput. Oper. Res.*, vol. 36, no. 7, pp. 2230–2241, Jul. 2009.
[8] H. Karimi and M. Bashiri, "Hub covering location problems with different coverage types," *Sci. Iranica*, vol. 18, no. 6, pp. 1571–1578, Dec. 2011.
[9] Y. H. Hwang and Y. H. Lee, "Uncapacitated single allocation p-hub maximal covering problem," *Comput. Ind. Eng.*, vol. 63, no. 2, pp. 382–389, Sep. 2012.
[10] M. Mohammadi, F. Jolai, and R. Tavakkoli-Moghaddam, "Solving a new stochastic multi-mode p -hub covering location problem considering risk by a novel multi-objective algorithm," *Appl. Math. Model.*, vol. 37, no. 24, pp. 10053–10073, Dec. 2013.
[11] H. Calık, S. A. Alumur, B. Y. Kara, and O. E. Karasan, "A tabu-search based heuristic for the hub covering problem over incomplete hub networks," *Comput. Oper. Res.*, vol. 36, no. 12, pp. 3088–3096, Dec. 2009.
[12] A. T. Ernst and M. Krishnamoorthy, "Efficient algorithms for the uncapacitated single allocation p-hub median problem," *Location Sci.*, vol. 4, no. 3, pp. 139–154, Oct. 1996.
[13] A. T. Ernst and M. Krishnamoorthy, "Solution algorithms for the capacitated single allocation hub location problem," *Ann. Oper. Res.*, vol. 86, pp. 141–159, 1999.
[14] A. Ilić, D. Urošević, J. Brimberg, and N. Mladenović, "A general variable neighborhood search for solving the uncapacitated single allocation p-hub median problem," *Eur. J. Oper. Res.*, vol. 206, no. 2, pp. 289–300, Oct. 2010.

[15] G. Li, D. Hu, and L. Su, "The model of location for single allocation multimodal hub under capacity constraints," *Procedia-Social Behav. Sci.*, vol. 96, pp. 351–359, Nov. 2013.
[16] A. T. Ernst and M. Krishnamoorthy, "Exact and heuristic algorithms for the uncapacitated multiple allocation p-hub median problem," *Eur. J. Oper. Res.*, vol. 104, no. 1, pp. 100–112, Jan. 1998.
[17] J. Ebery, M. Krishnamoorthy, A. Ernst, and N. Boland, "The capacitated multiple allocation hub location problem: Formulations and algorithms," *Eur. J. Oper. Res.*, vol. 120, no. 3, pp. 614–631, Feb. 2000.
[18] J. Kratica, "An electromagnetism-like metaheuristic for the uncapacitated multiple allocation p-hub median problem," *Comput. Ind. Eng.*, vol. 66, no. 4, pp. 1015–1024, Dec. 2013.
[19] N. Boland, M. Krishnamoorthy, A. T. Ernst, and J. Ebery, "Preprocessing and cutting for multiple allocation hub location problems," *Eur. J. Oper. Res.*, vol. 155, no. 3, pp. 638–653, Jun. 2004.
[20] G. Mayer and B. Wagner, "HubLocator: An exact solution method for the multiple allocation hub location problem," *Comput. Oper. Res.*, vol. 29, no. 6, pp. 715–739, May 2002.
[21] M. Mohammadi, P. Jula, and R. Tavakkoli-Moghaddam, "Design of a reliable multi-modal multi-commodity model for hazardous materials transportation under uncertainty," *Eur. J. Oper. Res.*, vol. 257, no. 3, pp. 792–809, Mar. 2017.
[22] H. Mohammadi, M. Ghazanfari, M. S. Pishvaee, and E. Teimoury, "Fresh-product supply chain coordination and waste reduction using a revenue-and-preservation-technology-investment-sharing contract: A real-life case study," *J. Cleaner Prod.*, vol. 213, pp. 262–282, Mar. 2019.
[23] M. Ghazanfari, H. Mohammadi, M. S. Pishvaee, and E. Teimoury, "Fresh-product trade management under government-backed incentives: A case study of fresh flower market," *IEEE Trans. Eng. Manag.*, vol. 66, no. 4, pp. 774–787, Nov. 2019.
[24] M. Fattahi, "A data-driven approach for supply chain network design under uncertainty with consideration of social concerns," *Ann. Oper. Res.*, vol. 97, pp. 31–47, Sep. 2020.
[25] M. Almiñana, L. F. Escudero, J. F. Monge, and J. Sánchez-Soriano, "On the enrouting protocol problem under uncertainty," *Eur. J. Oper. Res.*, vol. 181, no. 2, pp. 887–902, Sep. 2007.
[26] N. Gülpınar, D. Pachamanova, and E. Çanakoğlu, "Robust strategies for facility location under uncertainty," *Eur. J. Oper. Res.*, vol. 225, no. 1, pp. 21–35, Feb. 2013.
[27] M. Shahabi and A. Unnikrishnan, "Robust hub network design problem," *Transp. Res. E, Logistics Transp. Rev.*, vol. 70, pp. 356–373, Oct. 2014.
[28] Y. Li, Z. Yang, G. Li, D. Zhao, and W. Tian, "Optimal scheduling of an isolated microgrid with battery storage considering load and renewable generation uncertainties," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1565–1575, Feb. 2019.
[29] M. Majidi and K. Zare, "Integration of smart energy hubs in distribution networks under uncertainties and demand response concept," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 566–574, Jan. 2019.
[30] S. A. Alumur, S. Nickel, and F. Saldanha-da-Gama, "Hub location under uncertainty," *Transp. Res. B, Methodol.*, vol. 46, no. 4, pp. 529–543, May 2012.
[31] L. Zhen, "Storage allocation in transshipment hubs under uncertainties," *Int. J. Prod. Res.*, vol. 52, no. 1, pp. 72–88, Jan. 2014.
[32] Y. V. Hui, J. Gao, L. Leung, and S. Wallace, "Airfreight forwarder's shipment planning under uncertainty: A two-stage stochastic programming approach," *Transp. Res. E, Logistics Transp. Rev.*, vol. 66, pp. 83–102, Jun. 2014.
[33] T.-H. Yang, "Stochastic air freight hub location and flight routes planning," *Appl. Math. Model.*, vol. 33, no. 12, pp. 4424–4430, Dec. 2009.
[34] T.-H. Yang, "A two-stage stochastic model for airline network design with uncertain demand," *Transportmetrica*, vol. 6, no. 3, pp. 187–213, Jul. 2010.
[35] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, "Adjustable robust solutions of uncertain linear programs," *Math. Program.*, vol. 99, no. 2, pp. 351–376, Mar. 2004.
[36] A. Ben-Tal and A. Nemirovski, "Robust convex optimization," *Math. Oper. Res.*, vol. 23, no. 4, pp. 769–805, 1998.
[37] A. Ben-Tal and A. Nemirovski, "Robust solutions of linear programming problems contaminated with uncertain data," *Math. Program.*, vol. 88, no. 3, pp. 411–424, Sep. 2000.
[38] M. Meraklı and H. Yaman, "Robust intermodal hub location under polyhedral demand uncertainty," *Transp. Res. B, Methodol.*, vol. 86, pp. 66–85, Apr. 2016.
[39] E.-G. Talbi and R. Todosijević, "The robust uncapacitated multiple allocation p-hub median problem," *Comput. Ind. Eng.*, vol. 110, pp. 322–332, Aug. 2017.
[40] S. Mišković and Z. Stanimirović, "A hybrid metaheuristic method for the deterministic and robust uncapacitated multiple allocation p-hub centre problem," *Eur. J. Ind. Eng.*, vol. 11, no. 5, p. 631, 2017.
[41] S. Abbasi-Parizi, M. Aminnayeri, and M. Bashiri, "Robust solution for a minimax regret hub location problem in a fuzzy-stochastic environment," *J. Ind. Manage. Optim.*, vol. 14, no. 3, pp. 1271–1295, Jul. 2018.
[42] D. Bertsimas, D. Pachamanova, and M. Sim, "Robust linear optimization under general norms," *Oper. Res. Lett.*, vol. 32, no. 6, pp. 510–516, Nov. 2004.
[43] D. Bertsimas and M. Sim, "The price of robustness," *Oper. Res.*, vol. 52, no. 1, pp. 35–53, Feb. 2004.
[44] X. Chen, M. Sim, and P. Sun, "A robust optimization perspective on stochastic programming," *Oper. Res.*, vol. 55, no. 6, pp. 1058–1071, Dec. 2007.
[45] C. A. Zetina, I. Contreras, J.-F. Cordeau, and E. Nikbakhsh, "Robust uncapacitated hub location," *Transp. Res. B, Methodol.*, vol. 106, pp. 393–410, Dec. 2017.
[46] N. Ghaffari-Nasab, M. Ghazanfari, and E. Teimoury, "Robust optimization approach to the design of hub-and-spoke networks," *Int. J. Adv. Manuf. Technol.*, vol. 76, nos. 5–8, pp. 1091–1110, Feb. 2015.
[47] N. Ghaffarinasab, "An efficient matheuristic for the robust multiple allocation p-hub median problem under polyhedral demand uncertainty," *Comput. Oper. Res.*, vol. 97, pp. 31–47, Sep. 2018.
[48] D. Skorin-Kapov, J. Skorin-Kapov, and M. O'Kelly, "Tight linear programming relaxations of uncapacitated p-hub median problems," *Eur. J. Oper. Res.*, vol. 94, no. 3, pp. 582–593, Nov. 1996.

SHUXIA LI was born in Shiyan, Hubei, China, in 1975. She received the B.S. degree in mechanical engineering from the Hubei University of Automobile Technology, in 1996, the M.S. degree in mechanical engineering from Chongqing University, in 1999, and the Ph.D. degree in mechanical engineering from the Huazhong University of Science and Technology, in 2005. Since 2005, she has been an Assistant Professor and an Associate Professor with the School of Business, East China University of Science and Technology, Shanghai. Her research interests include modeling and optimization of logistics systems, scheduling theories and methods, and sustainable transportation.

CHENHAO FANG was born in Hangzhou, Zhejiang, China, in 1991. He received the B.S. degree in information and computing science from Zhejiang Sci-Tech University, in 2013, and the M.S. degree in management science and engineering from the East China University of Science and Technology, Shanghai, in 2016, where he is currently pursuing the Ph.D. degree with the School of Business. His research interests include design of logistics networks, robust optimization, and energy system modeling.

YIFAN WU was born in Wuhan, Hubei, China, in 1982. He received the B.S. and Ph.D. degrees in mechanical engineering from Shanghai Jiao Tong University, in 2004 and 2010, respectively. Since 2010, he was an Assistant Professor and an Associate Professor with the School of Business, East China University of Science and Technology, Shanghai. His research interests include design and optimization of logistics systems, simulation optimization, and supply chain management.