Hub location problem with balanced round-trip flows on hub links

Qing-Mi Hu

School of Economics and Management,
Jiangsu University of Science and Technology,
Zhenjiang 212003, China
Fax: +86-511-84401166
Email: huqingmi@just.edu.cn

Abstract: This paper addresses the classical single and multiple allocation hub location problems with fully interconnected hubs, in which the balanced round-trip flows on hub links are considered in the strategic decision making process. The use of balanced round-trip flows on hub links is motivated by the need to decrease the empty-trip rate of vehicles and increase the full-load rate of vehicles. Mixed-integer programming models are presented for single and multiple allocation versions of the problems. Numerical experiments with the CAB and AP datasets are performed to analyse the impacts of balanced flows on the network configurations and the utilisation of service resources. Experimental results show that the number of located hubs tends to decrease with a decrease in the allowable unbalanced round-trip degree of the hub link flows. The utilisation rate of transportation resources can be significantly improved with a small increase in the traditional operating cost. Moreover, a better modelling of economies of scale can be achieved when considering balanced flows. [Received: 22 September 2019; Accepted: 15 March 2020]

Keywords: hub location; balanced flows; economies of scale; mixed-integer programming.

Reference to this paper should be made as follows: Hu, Q-M. (2021) 'Hub location problem with balanced round-trip flows on hub links', *European J. Industrial Engineering*, Vol. 15, No. 1, pp.131–166.

Biographical notes: Qing-Mi Hu is a Lecturer in the School of Economics and Management, Jiangsu University of Science and Technology, China. His research interests include humanitarian logistics, second-order cone programming, and network design and optimisation.

1 Introduction

Hub-and-spoke networks (HSNs) have been extensively used in transportation, telecommunication, postal delivery, emergency service, and supply chain management (Berman et al., 2007; Cetiner et al., 2010; Ishfaq and Sox, 2012; Kim and O'Kelly, 2009; Lin and Lee, 2010). HSN is a common form of transportation networks, and used by most air transportation and logistics companies interested in exploiting economies of scale (O'Kelly and Bryan, 1998). HSNs can provide efficient transportation for origin-destination (OD) flows via a set of hubs that consolidate and switch the arrived flows, hub arcs (links) that connect two hubs, and spoke arcs that connect the non-hub nodes (spokes) and hubs. The crucial task in devising an efficient HSN is to determine the hub locations, hub capacities, and allocation of spokes to hubs. The goal is to minimise the operating cost, which typically includes the setup cost for hubs and the total routing cost for OD flows.

Hub location problems may be classified by the allocation pattern of spokes to hubs. One possibility is single allocation, in which each spoke is allocated to a single hub. A second possibility is multiple allocation, in which each spoke may be allocated to multiple hubs. In addition, there are two common assumptions in hub location problems. One is that no direct connection is possible between spokes. Another is that the hub network is a complete graph, which means that all hubs are directly connected. Nevertheless, in some applications, these assumptions are relaxed (Alumur et al., 2009; de Sá et al., 2018; Yoon and Current, 2008).

Studies of HSN design problems have attracted much attention since O'Kelly (1987). Subsequently, Campbell (1996) identified the differences between p-median and p-hub median problems, and proposed an integer linear program for the p-hub median problem with single and multiple allocations. Based on the studies of Campbell (1994, 1996), Ernst and Krishnamoorthy (1996, 1998), the formulations of HSNs are extended in the literature by incorporating different features such as networking policies (Aykin, 1995), hub capacities (Ebery et al., 2000), and hub as well as arc capacities (Sasaki and Fukushima, 2003). Recent research has addressed multiple capacity levels (Correia et al., 2010b), dynamic location (Contreras et al., 2011; Gelareh et al., 2015), competitive location (Ghaffarinasab et al., 2018; Lüer-Villagra and Marianov, 2013; Mahmutogullari and Kara, 2016; Sasaki et al., 2014), cooperative location (Cruijssen et al., 2010), multimodal hub location (Alumur et al., 2012), reliable location (An et al., 2015; Azizi et al., 2016), carrier collaboration (Hernandez et al., 2011), service considerations (Campbell, 2009; Chen et al., 2008; Lin and Lee, 2015), environment considerations (Hu, 2017; Mohammadi et al., 2014). For a comprehensive and early review of research on HSN design problems, refer to Alumur and Kara (2008), Campbell and O'Kelly (2012) and Farahani et al. (2013).

Recently, some studies have addressed the control of density of flows across the network by achieving the trade-off between minimisation of cost and congestion. For example, Elhedhli and Wu (2010) viewed the hub-and-spoke system as a network of M/M/1 queues, modelled congestion at hubs as the ratio of total flows to surplus capacities, and formulated a nonlinear mixed-integer programming model to explicitly minimise congestion, capacity acquisition, and transportation costs. de Camargo et al. (2011) dealt with a single allocation hub location problem under congestion, which is formulated as a nonlinear program and proposed a hybrid of outer approximation and Benders decomposition algorithm to solve it. Kian and Kargar (2016) presented the hub location problem with a power-law congestion cost and proposed two conic quadratic formulations to solve this problem. Azizi et al. (2018) established a nonlinear integer programming model to design HSNs under stochastic demand and congestion. They modelled hubs as spatially distributed M/G/1 queues and captured congestion using the expected queue lengths at hub facilities. Alkaabneh et al. (2019) considered a HSN design problem with inter-hub economies-of-scale and hub congestion. They explicitly modelled the economies-of-scale as a concave piece-wise linear function and congestion as a convex function, and a Lagrangian heuristic and a greedy randomised adaptive search procedure are applied to solve the nonlinear mixed-integer programming model.

Additionally, Monemi et al. (2020) controlled the density of flows across the HSN by achieving the trade-off between minimisation of cost and load balancing. They proposed a bi-objective nonlinear mixed-integer programming model to address the unbalanced spatial distribution of hub-level flows in HSNs from the infrastructure owner's point of view.

Although addressing the control of density of flows across the HSN has received more attention from the perspective of minimising hub congestion effects, studies concerning the balanced distribution of flows across the HSN structure are rare. As Monemi et al. (2020) pointed out, from the infrastructure owner perspective, a balanced distribution of flows across the HSN is a critical issue, which should be carefully considered in HSNs design. Meanwhile, studies addressing the balanced round-trip flows on hub links appear to be lacking.

Actually, in HSN operations, hubs not only serve as the locations where flows are consolidated and switched, but also act as the home bases for vehicles or drivers. A variety of vehicles can be used in HSNs. Smaller vehicles (or airplanes) are used to transfer flows between spokes and hubs, while larger and more efficient vehicles handle the transfers between hubs. However, the routed OD flows in HSNs are usually unbalanced in opposing directions (round-trip) on a hub link in practice. These unbalanced round-trip flows may increase transportation costs, since there is considerable increase in the empty-trip rate of vehicles and a decrease in the full-load rate of vehicles. Thus, to avoid travelling empty, some drivers may spend more time seeking loads, which greatly decreases the efficiency of vehicles or drivers and increases the away-from-home times for drivers. In fact, with proper strategic design of the network, empty travel can be better controlled in HSNs. For example, by imposing the balancing requirements of the round-trip flows in the hub links on the network design, the utilisation of transportation resources (e.g., vehicles and drivers) can be significantly improved, and regular schedules for vehicles can be implemented. Therefore, based on the fact that vehicles depart from one hub and arrive at another hub, and then must return to the hub (or home base) in HSN operations, it makes sense to balance the round-trip flows routed via the hub links in logistics and transport industry.

To illustrate the impact of balanced round-trip flows in the hub links on HSN design and resource utilisation, a small example is presented. There are ten nodes scattered in a Euclidean plane as shown in Figure 1. It is assumed that:

- The distance between two nodes in HSN is given by the Euclidean distance.
- Each node (except for nodes 1 and 8) sends two units of flow to every other node but not to itself. Nodes 1 and 8 send eight units of flow to every other node but not to itself. Accordingly, each node (except for nodes 1 and 8) originates 18 units of flow. Nodes 1 and 8 originate 72 units of flow respectively.
- A hub can be established in every node. Two hubs are considered to be established. The fixed setup costs for the hubs are equal to 200 monetary units.
- The cost for transferring one unit of flow between two hubs is equal to 0.6 times the distance between the hubs. The cost for transferring one unit of flow between any pair of non-hub nodes and hub nodes is equal to the distance between the nodes involved.

The optimal HSN for the illustrated instance is given in Figure 2, where the squared nodes represent the established hubs. In the optimal solution, nodes 4 and 5 are chosen as hubs. Nodes 1, 2, 6, 8 and 9 are allocated to hub 4 and nodes 3, 7 and 10 are allocated to hub 5. The total operating cost is equal to 1,527.5 monetary units. Additionally, it can be seen that the total flows routed from hub 4 to hub 5 are equal to 96 units. However, the total flows routed from hub 5 to hub 4 are significantly low and only reach 48 units.

The solution depicted in Figure 2 is clearly unbalanced in terms of the transferred flows in the round-trip of the hub link transportation. Suppose that a balancing requirement is imposed such that the absolute value of the difference between the round-trip flows routed via the hub link should be at most equal to 10% of the total flows routed via the hub link. In this case, the optimal HSN significantly changes. Figure 3 depicts the new solution, where nodes 4 and 6 are chosen as hubs. The total operating cost for the new solution is equal to 1,538.5 monetary units. Moreover, the round-trip flows are balanced in the hub link transportation. The total flows routed from hub 4 to hub 6 are equal to the total flows routed from hub 6 to hub 4.

We define the unbalanced round-trip degree of a hub link as the ratio of the absolute value of the difference between the round-trip flows routed via the hub link to the total flows routed via the hub link. Comparison of the solutions for the unbalanced and balanced flows (Figures 2 and 3) shows that the unbalanced round-trip degree of the hub link can be greatly decreased with a small increase in the total operating cost. For example, the unbalanced round-trip degree of the hub link can be reduced from 0.3333 to 0, or 100%, with a small increase in total operating cost from 1,527.5 to 1,538.5 (0.72%). Furthermore, if the load capacities of a vehicle (or airplane) are equal to eight units of flow, ten vehicles are required in the round-trip of the hub link transportation for balanced flows (Figure 3). The empty-trip rate of the vehicles is then zero and the full-load rate of vehicles reaches 100%. In other words, all vehicles are fully utilised in the departure and return trips. However, for unbalanced flows (Figure 2), 12 vehicles are required in the round-trip of the hub link transportation. The empty-trip rate of vehicles increases to 50% and the full-load rate of vehicles decreases to 50% in the return trip, since six empty vehicles still need to return from hub 5 to hub 4. In this case, the high empty-trip rate of vehicles and the low full-load rate of vehicles significantly decrease the utilisation of transportation resources and increase the transportation cost. In particular, the increase in the transportation cost caused by unbalanced flows is obvious in the airline industry because air transportation is more expensive than other modes of transportation. Additionally, although the choice of the capacity of the vehicle that better fits the amount of flow to be transported may contribute to the optimisation of costs, there is no guarantee that these chosen vehicles are fully utilised in their return trip. Therefore, to optimise resource utilisation, it is vital that the balanced round-trip flows on hub links should be carefully considered in HSN design and operations.

To the best of the authors' knowledge, only Kewcharoenwong and Üster (2017) addressed the control of unbalanced round-trip flows on links in the context of logistics network design problems. To improve operational efficiency via increased ability to return drivers to their home bases and reduce empty backhauls, Kewcharoenwong and Üster (2017) introduced link capacity constraints and the concept of link imbalance in the strategic relay network design, developed a mixed-integer programming model and devised an efficient Lagrangean decomposition algorithm to solve large-size problems.

Figure 1 A set of nodes of an illustrated example

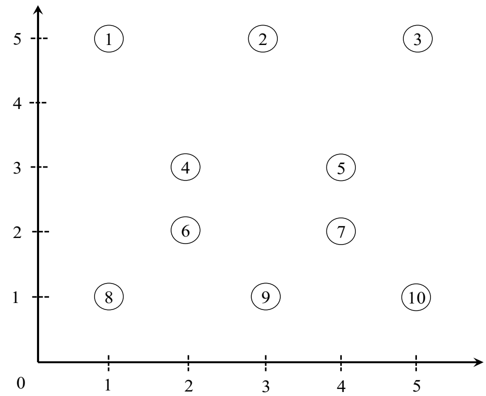


Figure 2 Optimal HSN with unbalanced flows

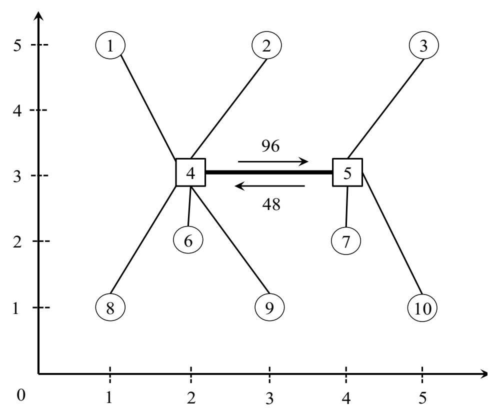
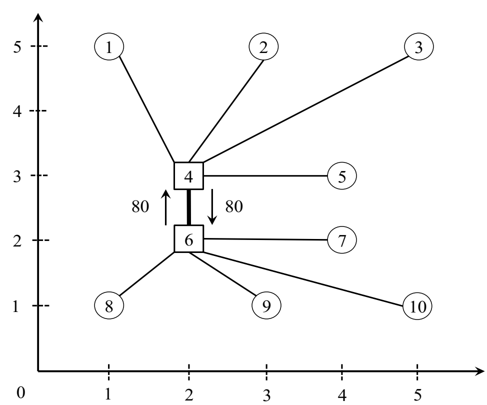


Figure 3 Optimal HSN with balanced flows



This paper focuses on the classical single and multiple allocation hub location problems with fully interconnected hubs, in which the balanced round-trip flows on hub links are incorporated into HSN design to optimise resource utilisation. Moreover, the impacts of balanced flows on HSN configurations are analysed, and the benefits from balanced flows are investigated. To our knowledge, this study is first to address balanced flows on hub links in the strategic HSN design.

The remainder of this paper is organised as follows: In Section 2, mixed-integer programming models for single and multiple allocation cases are formulated. In Section 3, computational experiments are performed and the experimental results are reported. The work is concluded and further research directions are suggested in Section 4.

2 Formulation

This paper considers uncapacitated hub location problems under the assumptions that the subgraph induced by the hubs is complete, and there are no direct links between spokes.

2.1 Notation

The following notation is used in Subsections 2.2 and 2.3 to present the formulation.

Sets	
N	A set of nodes: $N = \{1, 2, \dots, n\}$, and indexed by i, j, k, l
Parameters	
f_k	Fixed setup cost for establishing a hub at node k
$d_{i,j}$	Distance from node i to node j, $d_{i,i} = 0$, and the distances satisfy the triangle inequality
$w_{i,j}$	Flows originated at node i and destined to node j
O_i	Total flows originated at node i, $O_i = \sum_j w_{i,j}$
D_i	Total flows destined to node i, $D_i = \sum_j w_{j,i}$
χ	Collection cost per unit of flow and per unit of distance between a spoke and a hub
δ	Distribution cost per unit of flow and per unit of distance between a hub and a spoke
α	Transfer cost per unit of flow and per unit of distance between hubs, it is assumed that $\alpha < \chi$ and $\alpha < \delta$
θ	Allowable unbalanced round-trip degree of the hub link flows

2.2 Balanced flows

The goal of balanced round-trip flows on hub links is to balance and improve the utilisation of transportation resources in the round-trip of the hub link transportation. For example, the empty-trip rate of vehicles is decreased and the full-load rate of vehicles is increased in the round-trip. Since the round-trip flows on hub links relate to the hub location decision and the allocation of spokes to hubs, balancing the round-trip flows on hub links should be solved in the strategic decision making process. Once the round-trip flows on hub links are balanced and determined, the corresponding load capacity and number of vehicles can be better chosen to fit to the routed flows. Meanwhile, the chosen vehicles also can be fully utilised in the departure and return trips.

Ernst and Krishnamoorthy (1996) proposed a set of three-index flow variables for formulating hub location problems, which led to models that could better achieve a trade-off between compactness and tightness. In this paper, we adopt such flow variables for both the single and multiple allocation versions:

$y_{k,l}^{i}$	Amount of flow originated at node *i* that is routed from hub *k* to *l*

Further, using the three-index flow variables $y_{k,l}^{i}$, we consider several methods to achieve the balanced round-trip flows on hub links. First, by penalising the unbalanced round-trip flows, an objective function *F* is denoted by equation (1) to minimise the total penalty cost, where ϕ is a coefficient of cost for penalising a unit of unbalanced round-trip flow, and $\sum_i y_{k,l}^{i}$ represents the total flows routed from hub *k* to *l*. Similarly, $\sum_i y_{l,k}^{i}$ represents the total flows routed from hub *l* to *k*.

$$F = \sum_{k,l,k<l} \phi \cdot \left| \sum_i y_{k,l}^{i} - \sum_i y_{l,k}^{i} \right| \tag{1}$$

Second, by imposing the restriction that unbalanced round-trip flows are smaller than or equal to a predefined value ($\Im$), the balanced round-trip flow constraints (BRFCs) are denoted by inequalities (2), where $\Im \geq 0$.

$$\left| \sum_i y_{k,l}^{i} - \sum_i y_{l,k}^{i} \right| \leq \Im, \qquad \forall k, l, k < l \tag{2}$$

However, the objective function (1) is no guarantee that the round-trip flows on hub links are balanced. Moreover, since it is still unclear what flow amounts are transferred via the hub link, it is difficult to determine an appropriate parameter for $\Im$ and evaluate the rationality of the parameter of $\Im$, since a looser value of $\Im$ may result in an unsatisfied solution and a tighter value of $\Im$ may result in an infeasible solution. To overcome these limitations, the BRFCs are redefined as inequalities (3). Inequalities (3) ensure that the absolute value of the difference between the flows routed from hub *k* to *l* and from hub *l* to *k* should be less than or equal to θ times the total flows routed via the hub link (*k*, *l*).

$$\left| \sum_i y_{k,l}^{i} - \sum_i y_{l,k}^{i} \right| \leq \theta \cdot \left(\sum_i y_{k,l}^{i} + \sum_i y_{l,k}^{i} \right), \qquad \forall k, l, k < l \tag{3}$$

2.3 Model

In this section, two mixed-integer programming models are built by imposing the balanced round-trip flows in the hub links on the classical single and multiple allocation hub location problems.

Based on the formulations proposed by Ernst and Krishnamoorthy (1999) and Correia et al. (2010a), the single allocation hub location problem with balanced flows, model [M1], is formulated as follows:

$$[\text{M1}] \quad \min F = \sum_k f_k \cdot x_{k,k} + \sum_i \sum_k d_{i,k} \cdot (\chi \cdot O_i + \delta \cdot D_i) \cdot x_{i,k} + \sum_i \sum_k \sum_l \alpha \cdot d_{k,l} \cdot y^i_{k,l} \quad (4)$$

s.t.

$$\sum_k x_{i,k} = 1, \quad \forall i \quad (5)$$

$$x_{i,k} \le x_{k,k}, \quad \forall i, k \quad (6)$$

$$\sum_l y^i_{k,l} - \sum_l y^i_{l,k} = O_i \cdot x_{i,k} - \sum_j w_{i,j} \cdot x_{j,k}, \quad \forall i, k \quad (7)$$

$$\sum_l y^i_{k,l} \le O_i \cdot x_{i,k}, \quad \forall i, k \quad (8)$$

$$\left| \sum_i y^i_{k,l} - \sum_i y^i_{l,k} \right| \le \theta \cdot \left(\sum_i y^i_{k,l} + \sum_i y^i_{l,k} \right), \quad \forall k, l, k < l \quad (9)$$

$$x_{i,k} \in \{0, 1\}, \quad \forall i, k \quad (10)$$

$$y^i_{k,l} \ge 0, \quad \forall i, k, l \quad (11)$$

In model [M1], $x_{i,k}$ is a binary variable that is equal to 1 if node i is allocated to hub k and 0 otherwise.

The objective function (4) minimises the total operating cost, which consists of three components:

1 the setup cost of hubs

2 the collection cost of the flows from spokes to hubs and the distribution cost of the flows from hubs to spokes

3 the transfer cost of the flows between hubs.

Constraint (5) ensures that each node is a hub or is allocated to a hub. Constraint (6) ensures that a spoke is only allocated to an operating hub. Constraint (7) is a divergence equation at node k for the flows associated with node i. Constraint (8) ensures that a $y^i_{k,l}$ can only be different from 0 if $x_{i,k}$ is equal to one and in this case, all the flows originating at node i are routed to hub k. Constraint (9), the BRFC, has the same meaning as inequalities (3). Constraints (10) and (11) are domain constraints.

Based on the formulation proposed by Ernst and Krishnamoorthy (1998), the multiple allocation hub location problem with balanced flows, model [M2], is formulated as follows:

$$[\text{M2}] \quad \min F = \sum_k f_k \cdot x_k + \sum_i \sum_k \chi \cdot d_{i,k} \cdot u_{i,k} + \sum_i \sum_k \sum_l \alpha \cdot d_{k,l} \cdot y^i_{k,l} + \sum_i \sum_l \sum_j \delta \cdot d_{l,j} \cdot v^i_{l,j} \quad (12)$$

s.t.

$$\sum_k u_{i,k} = O_i, \qquad \forall i \qquad (13)$$

$$\sum_l v_{l,j}^i = w_{i,j}, \qquad \forall i, j \qquad (14)$$

$$\sum_l y_{k,l}^i - \sum_l y_{l,k}^i = u_{i,k} - \sum_j v_{k,j}^i, \qquad \forall i, k \qquad (15)$$

$$\left| \sum_i y_{k,l}^i - \sum_i y_{l,k}^i \right| \le \theta \cdot \left(\sum_i y_{k,l}^i + \sum_i y_{l,k}^i \right), \qquad \forall k, l, k < l \qquad (16)$$

$$\sum_l y_{k,l}^i \le u_{i,k}, \qquad \forall i, k \qquad (17)$$

$$u_{i,k} \le O_i \cdot x_k, \qquad \forall i, k \qquad (18)$$

$$v_{l,j}^i \le w_{i,j} \cdot x_l, \qquad \forall i, j, l \qquad (19)$$

$$x_k \in \{0, 1\}, \qquad \forall k \qquad (20)$$

$$y_{k,l}^i, u_{i,k}, v_{l,j}^i \ge 0, \qquad \forall i, j, k, l \qquad (21)$$

In model [M2], x_k is a binary variable that is equal to 1 if node k is selected to be a hub and 0 otherwise; $u_{i,k}$ represents the total flows that are sent from node i directly to hub k; $v_{l,j}^i$ represents the flows originated at node i that flow from hub l to node j.

The objective function (12) minimises the total operating cost including the setup cost of hubs, the collection, transfer and distribution costs of the flows. Constraint (13) ensures that all flows are collected at the hubs. Constraint (14) ensures that all flows are distributed from the hubs. Constraint (15) is a flow divergence constraint. Constraint (16) is the BRFC. Constraints (17)–(19) are consistency limitations that make the definition of the variables meaningful. Constraints (20) and (21) are domain constraints.

Models [M1] and [M2] show that the BRFCs are nonlinear and can be linearised as constraints (22) and (23). Alternatively, the BRFCs also can be linearised as constraints (24) and (25) by introducing an auxiliary variable $\Gamma_{k,l}$.

$$(1-\theta) \cdot \sum_i y_{k,l}^i \le (1+\theta) \cdot \sum_i y_{l,k}^i, \qquad \forall k, l, k < l \qquad (22)$$

$$(1-\theta) \cdot \sum_i y_{l,k}^i \le (1+\theta) \cdot \sum_i y_{k,l}^i, \qquad \forall k, l, k < l \qquad (23)$$

$$\Gamma_{k,l} + (1-\theta) \cdot \sum_i y_{k,l}^i = (1+\theta) \cdot \sum_i y_{l,k}^i, \qquad \forall k, l, k < l \qquad (24)$$

$$0 \le \Gamma_{k,l} \le 2\theta \cdot \left(\sum_i y_{k,l}^i + \sum_i y_{l,k}^i \right), \qquad \forall k, l, k < l \qquad (25)$$

Proposition 1: If $\theta^* \leq \theta$, then the optimal value of model [M1] (or model [M2]) with the value of θ^* is greater than or equal to that of model [M1] (or model [M2]) with the value of θ.

Proof: When $\theta^* \leq \theta$, the optimal solution obtained based on the value of θ^* is a feasible solution for the case of the value of θ. Therefore, we can derive Proposition 1.

□

Remark 1: If $\theta \geq \max_{k,l,k<l} \{\vartheta_{k,l}\}$, then the BRFCs in model [M1] are not binding, where $\vartheta_{k,l}$ represents the unbalanced round-trip degree of the hub link (k, l) and is determined by equation (26), based on the solution of the model with (4)–(8), and (10)–(11). Similarly, if $\theta \geq \max_{k,l,k<l} \{\vartheta_{k,l}\}$, then the BRFCs in model [M2] are not binding, where $\vartheta_{k,l}$ is determined by equation (26), based on the solution of the model with (12)–(15), and (17)–(21).

$$\vartheta_{k,l} = \frac{\left| \sum_i y_{k,l}^i - \sum_i y_{l,k}^i \right|}{\sum_i y_{k,l}^i + \sum_i y_{l,k}^i}, \quad \forall k, l, k < l \tag{26}$$

Remark 2: Given any value of θ ($\theta \geq 0$), models [M1] and [M2] are always feasible, since the solution with only one established hub is always a feasible solution for the case of any θ ($\theta \geq 0$) value.

3 Numerical experiments

Numerical experiments are conducted to calculate the results of applying the [M1] and [M2] models to simulation cases. The models are formulated in a MATLAB environment, and solved by an IBM ILOG CPLEX 12.6. All numerical experiments are performed on a personal computer with an Intel Core i5 2.67 GHz CPU and 4 GB RAM.

3.1 Input parameters

In the following section, the performance of the proposed models is tested and the effects of balanced flows on optimal solutions are analysed based on the well-known CAB and AP datasets (O'Kelly, 1987; Ernst and Krishnamoorthy, 1996). The CAB dataset has symmetric demand and distance matrices between every pair of nodes. The AP dataset has an asymmetric demand matrix, coordinates for each node to compute the Euclidean distance between every pair of nodes, and hub fixed costs.

We denote $w_{i,j}^*$ and $d_{i,j}^*$ as the demand and distance between node i and node j in the CAB dataset introduced by O'Kelly (1987), respectively. In our experiments involving the CAB dataset, we set $d_{i,j} = d_{i,j}^*$ $(\forall i, j)$, $w_{i,j} = w_{i,j}^*$ $(\forall i \geq j)$. To analyse the impacts of asymmetric OD flows on network structures under the consideration of balanced flows, we specifically set $w_{i,j} = \lambda \cdot w_{i,j}^*$ $(\forall i < j)$, where λ is defined as the asymmetric degree of OD flows, and $\lambda \geq 1$. We let $\lambda \in \{1, 2, 3, 5, 7, 10\}$. Note that the case of $\lambda = 1$ represents the symmetric OD flows. Moreover, several related parameters are set as follows: To keep the hub fixed setup cost within the same order of magnitude of the transportation cost, f_k is set to 4,500 O_k ($\forall k$) because the CAB dataset does not have hub fixed costs. In our experiments involving the AP dataset, we choose the tight fixed setup cost for hubs. These data can be obtained from OR-Library (2015). In addition, since the total OD flows increase with increase in the value of λ in our experiments, we define $\overline{F}$ as the operating cost of unit flow to analyse the effects of asymmetric OD flows on the objective function value, where $\overline{F} = F \Big/ \sum_{i,j} w_{i,j}$.

To investigate the impacts of allowable unbalanced round-trip degree of the hub link flows on optimal solutions, we set $\theta \in \{0, 0.01, 0.04, 0.07, 0.1, 1\}$. Note that the case of $\theta = 0$ represents completely balanced round-trip flows on hub links, and the case of $\theta = 1$ represents that balanced flows are not considered. Finally, in our all experiments, the values of χ and δ are set to 1. The economy of scale factor is taken within the set $\alpha \in \{0.2, 0.4, 0.6, 0.8\}$, as is customary in the literature.

3.2 Analysing network configurations

To analyse how the parameters of α, θ and λ affect the hub network configurations, we solve different single and multiple allocation instances by varying these parameters. Numerical results are reported in Tables 1–4 and Figures 4–13.

From the results presented in Tables 1–4, we can clearly observe that the operating cost of unit flow increases with increase in economy of scale factor α for all single and multiple allocation instances in given the values of θ and λ. This trend, which is also valid in hub location problems with unbalanced flows, can be attributed to the fact that a higher discount can be achieved by decreasing the value of α for the aggregated flows. Moreover, as we know, increasing the value of α in general leads to fewer hubs for the classical hub location problems. This phenomenon can also be observed in the multiple allocation case with balanced flows (Tables 3–4). Because of the flexibility of the multiple allocation, the feasible solutions satisfying the BRFCs can be easily achieved regardless of the number of located hubs, which makes the optimal solutions more sensitive to the economy of scale factor than the value of θ. Therefore, greater scale economies (e.g., α = 0.2) encourage the location of more hubs so that more scale economies can be exploited. When lower scale economies are perceived (e.g., α = 0.8), then there is little incentive to locate more hubs to attain lower transportation costs. However, for the single allocation case with balanced flows, the number of located hubs appears to be insensitive to changes in economy of scale factor α (Tables 1–2). In particular, the hub location or network structure seems to be same across different economy of scale factors in the case of having low θ value. For example, when θ is set to 0, the same hub network is configured under different economy of scale factors. This is because the BRFCs restrict the feasible space of model [M1], which makes the optimal solutions more sensitive to θ value compared to the economy of scale factor. Because of the inflexibility of the single allocation, the number of located hubs significantly affects the feasibility and optimality of model [M1]. For example, with a given θ value, locating more hubs may lead to an infeasible solution while the solution with fewer hubs may not be optimal under the BRFCs. Therefore, the number of located hubs has a low sensitivity to the economy of scale factor and largely depends on the value of θ.

Table 1 Single allocation results for CAB dataset instances with 25 nodes

λ	θ	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
		$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*
1	1	1,049	2, 5, 13, 19, 24	1,182	2, 5, 13, 19	1,299	2, 5, 19	1,409	2, 5, 19
	0.1	1,049	2, 5, 13, 19, 24	1,182	2, 5, 13, 19	1,299	2, 5, 19	1,409	2, 5, 19
	0.07	1,049	2, 5, 13, 19, 24	1,182	2, 5, 13, 19	1,299	2, 5, 19	1,409	2, 5, 19
	0.04	1,049	2, 5, 13, 19, 24	1,182	2, 5, 13, 19	1,299	2, 5, 19	1,409	2, 5, 19
	0.01	1,049	2, 5, 13, 19, 24	1,182	2, 5, 13, 19	1,299	2, 5, 19	1,409	2, 5, 19
	0	1,049	2, 5, 13, 19, 24	1,182	2, 5, 13, 19	1,299	2, 5, 19	1,409	2, 5, 19
2	1	1,042	2, 5, 13, 19, 24	1,176	2, 5, 13, 19, 24	1,309	2, 5, 13, 19, 24	1,421	13, 19, 20
	0.1	1,103	2, 13, 19, 24	1,231	13, 19, 20, 24	1,339	13, 19, 20	1,446	13, 19, 20
	0.07	1,139	2, 13, 19	1,249	13, 19, 20	1,352	13, 19, 20	1,454	13, 19, 20
	0.04	1,174	2, 19, 21	1,299	13, 19, 20	1,409	19, 20	1,512	19, 20
	0.01	1,190	2, 19, 21	1,321	2, 19, 21	1,438	19, 20	1,536	19, 20
	0	1,572	5	1,572	5	1,572	5	1,572	5
3	1	1,018	18, 19, 21, 24	1,156	13, 19, 20, 24	1,278	13, 19, 20, 24	1,397	13, 19, 20, 24
	0.1	1,112	19, 20, 24	1,225	13, 19, 20	1,328	13, 19, 20	1,431	13, 19, 20
	0.07	1,150	13, 19, 20	1,271	13, 19, 20	1,376	19, 20	1,474	19, 20
	0.04	1,162	2, 19, 21	1,291	13, 19, 20	1,408	13, 19, 20	1,506	13, 19, 20
	0.01	1,193	5, 18, 19	1,318	19, 20	1,416	19, 20	1,514	19, 20
	0	1,578	5	1,578	5	1,578	5	1,578	5
5	1	964	18, 19, 21, 23, 24	1,109	19, 21, 24, 25	1,233	19, 20, 21, 24	1,351	19, 20, 21, 24
	0.1	1,106	19, 20, 24	1,238	19, 20, 21	1,344	19, 20, 21	1,450	19, 20, 21
	0.07	1,128	13, 19, 20	1,250	19, 21, 25	1,364	19, 20	1,467	19, 20
	0.04	1,135	18, 19, 21	1,252	19, 21, 25	1,380	19, 21, 25	1,479	20, 21

Table 1 Single allocation results for CAB dataset instances with 25 nodes (continued)

λ	θ	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
		$\overline{F}$	*Hub location*	$\overline{F}$	*Hub location*	$\overline{F}$	*Hub location*	$\overline{F}$	*Hub location*
5	0.01	1,197	5, 18, 19	1,296	19, 20	1,394	19, 20	1,492	19, 20
	0	1,577	20	1,577	20	1,577	20	1,577	20
7	1	923	19, 21, 23, 24, 25	1,064	19, 21, 23, 24, 25	1,197	19, 21, 23, 24, 25	1,325	19, 20, 21, 24
	0.1	1,081	19, 21, 24, 25	1,217	19, 21, 25	1,342	19, 21, 25	1,437	19, 20, 21
	0.07	1,088	19, 21, 25	1,217	19, 21, 25	1,343	19, 21, 25	1,455	20, 21
	0.04	1,090	19, 21, 25	1,219	19, 21, 25	1,347	19, 21, 25	1,463	20, 21
	0.01	1,143	19, 21, 25	1,270	19, 21, 25	1,382	19, 20	1,475	20, 21
	0	1,569	20	1,569	20	1,569	20	1,569	20
10	1	885	19, 21, 23, 24, 25	1,026	19, 21, 23, 24, 25	1,159	19, 21, 23, 24, 25	1,293	19, 21, 23, 24, 25
	0.1	1,061	19, 21, 25	1,190	19, 21, 25	1,315	19, 21, 25	1,421	19, 20, 21
	0.07	1,061	19, 21, 25	1,190	19, 21, 25	1,318	19, 21, 25	1,447	19, 20
	0.04	1,063	19, 21, 25	1,191	19, 21, 25	1,320	19, 21, 25	1,448	19, 21, 25
	0.01	1,116	19, 21, 25	1,243	19, 21, 25	1,370	19, 21, 25	1,454	21, 25
	0	1,563	20	1,563	20	1,563	20	1,563	20

Table 2 Single allocation results for AP dataset instances with 20 and 25 nodes

n	θ	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
		$\overline{F}$	*Hub location*	$\overline{F}$	*Hub location*	$\overline{F}$	*Hub location*	$\overline{F}$	*Hub location*
20	1	119	1, 8, 10, 16, 19, 20	148	1, 8, 10, 16, 17, 20	176	1, 8, 10, 11, 17, 20	204	1, 8, 10, 11, 17, 20
	0.1	146	7, 8, 16, 19, 20	170	7, 8, 16, 17, 20	194	7, 8, 16, 17, 20	218	7, 8, 16, 17, 20
	0.07	181	7, 11, 19	200	7, 16, 17	220	7, 16, 17	235	7, 17, 20
	0.04	199	7, 17, 18	212	7, 17, 20	223	7, 17, 20	235	7, 17, 20

Table 2 Single allocation results for AP dataset instances with 20 and 25 nodes (continued)

n	θ	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
		$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*
20	0.01	228	7, 20	241	7, 20	253	7, 17	259	10
	0	259	10	259	10	259	10	259	10
25	1	174	1, 5, 10, 16, 19, 21, 24	210	1, 5, 10, 13, 16, 21, 24	242	1, 5, 10, 14, 21, 24	275	1, 5, 10, 14, 21, 24
	0.1	199	1, 5, 10, 14, 21	233	1, 5, 10, 14, 21	264	5, 10, 14, 21	285	10, 21, 24
	0.07	207	4, 5, 10, 14, 21	239	5, 10, 14, 21	265	10, 14, 21	287	10, 14, 21
	0.04	223	4, 5, 14, 15	252	7, 14, 15	277	10, 14, 15	297	10, 21, 24
	0.01	241	10, 14	262	10, 14	283	10, 14	305	10, 14
	0	355	10	355	10	355	10	355	10

Table 3 Multiple allocation results for CAB dataset instances with 25 nodes

λ	θ	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
		$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*
1	1	1,046	2, 5, 13, 19, 24	1,145	2, 5, 13, 19	1,216	2, 13, 19	1,273	2, 13, 19
	0.1	1,046	2, 5, 13, 19, 24	1,145	2, 5, 13, 19	1,216	2, 13, 19	1,273	2, 13, 19
	0.07	1,046	2, 5, 13, 19, 24	1,145	2, 5, 13, 19	1,216	2, 13, 19	1,273	2, 13, 19
	0.04	1,046	2, 5, 13, 19, 24	1,145	2, 5, 13, 19	1,216	2, 13, 19	1,273	2, 13, 19
	0.01	1,046	2, 5, 13, 19, 24	1,145	2, 5, 13, 19	1,216	2, 13, 19	1,273	2, 13, 19
	0	1,046	2, 5, 13, 19, 24	1,145	2, 5, 13, 19	1,216	2, 13, 19	1,273	2, 13, 19
2	1	1,038	2, 5, 13, 19, 24	1,144	2, 13, 19, 24	1,222.8	2, 13, 19	1,279.8	2, 13, 19
	0.1	1,040	2, 19, 21, 24	1,145	2, 13, 19, 24	1,223	2, 13, 19	1,279.8	2, 13, 19
	0.07	1,041	2, 19, 21, 24	1,145.5	2, 13, 19, 24	1,223.2	2, 13, 19	1,279.9	2, 13, 19
	0.04	1,042	2, 19, 21, 24	1,146	2, 13, 19, 24	1,223.6	2, 13, 19	1,279.9	2, 13, 19
	0.01	1,043	2, 19, 21, 24	1,147.5	2, 13, 19, 24	1,224	2, 13, 19	1,280	2, 13, 19
	0	1,043	2, 19, 21, 24	1,148	2, 13, 19, 24	1,224.2	2, 13, 19	1,280.1	2, 13, 19

Table 3 Multiple allocation results for CAB dataset instances with 25 nodes (continued)

λ	θ	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
		$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*
3	1	1,005.3	18, 19, 21, 24	1,119.6	18, 19, 21, 24	1,214	13, 19, 20, 24	1,273	13, 19, 20
	0.1	1,010.7	18, 19, 21, 24	1,125.2	18, 19, 21, 24	1,215.7	13, 19, 20, 24	1,273.4	13, 19, 20
	0.07	1,012.1	18, 19, 21, 24	1,126.6	18, 19, 21, 24	1,216.2	13, 19, 20, 24	1,273.5	13, 19, 20
	0.04	1,013.6	18, 19, 21, 24	1,128.2	18, 19, 21, 24	1,216.9	13, 19, 20, 24	1,273.6	13, 19, 20
	0.01	1,016	18, 19, 21, 24	1,130	18, 19, 21, 24	1,217.9	13, 19, 20, 24	1,273.7	13, 19, 20
	0	1,017	18, 19, 21, 24	1,130.7	18, 19, 21, 24	1,218.4	13, 19, 20, 24	1,273.8	13, 19, 20
5	1	949.8	18, 19, 21, 23, 24	1,068.3	18, 19, 21, 23, 24	1,162.5	19, 21, 24, 25	1,240.2	19, 21, 24, 25
	0.1	962.8	18, 19, 21, 23, 24	1,080.1	19, 21, 24, 25	1,168.3	19, 21, 24, 25	1,243.3	19, 21, 24, 25
	0.07	965.7	18, 19, 21, 23, 24	1,081.7	19, 21, 24, 25	1,169.8	19, 21, 24, 25	1,243.7	19, 21, 24, 25
	0.04	970.4	18, 19, 21, 23, 24	1,083.3	19, 21, 24, 25	1,171.2	19, 21, 24, 25	1,244.3	19, 21, 24, 25
	0.01	975.8	18, 19, 21, 23, 24	1,085.3	19, 21, 24, 25	1,172.6	19, 21, 24, 25	1,245	19, 20, 21, 24
	0	977.3	18, 19, 21, 24	1,086	19, 21, 24, 25	1,173.1	19, 21, 24, 25	1,245.1	19, 20, 21, 24
7	1	911.97	19, 21, 23, 24, 25	1,022.5	19, 21, 23, 24, 25	1,120.9	19, 21, 23, 24, 25	1,200.6	19, 21, 24, 25
	0.1	927	19, 21, 23, 24, 25	1,039.6	19, 21, 23, 24, 25	1,130.8	19, 21, 24, 25	1,204.3	19, 21, 24, 25
	0.07	931.7	19, 21, 23, 24, 25	1,042.2	19, 21, 23, 24, 25	1,132.3	19, 21, 24, 25	1,204.9	19, 21, 24, 25
	0.04	936.8	19, 21, 23, 24, 25	1,045	19, 21, 23, 24, 25	1,133.7	19, 21, 24, 25	1,205.6	19, 21, 24, 25
	0.01	942	19, 21, 23, 24, 25	1,048.5	19, 21, 23, 24, 25	1,135.2	19, 21, 24, 25	1,206.3	19, 21, 24, 25

Table 3 Multiple allocation results for CAB dataset instances with 25 nodes (continued)

λ	θ	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
		$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*
7	0	943.6	19, 21, 23, 24, 25	1,049.4	19, 21, 24, 25	1,135.8	19, 21, 24, 25	1,206.5	19, 21, 24, 25
10	1	874	19, 21, 23, 24, 25	984.6	19, 21, 23, 24, 25	1,083	19, 21, 23, 24, 25	1,166.5	19, 21, 23, 24, 25
	0.1	894.2	19, 21, 23, 24, 25	1,005	19, 21, 23, 24, 25	1,097.8	19, 21, 23, 24, 25	1,172.5	19, 21, 24, 25
	0.07	899	19, 21, 23, 24, 25	1,007.8	19, 21, 23, 24, 25	1,099.5	19, 21, 23, 24, 25	1,173.2	19, 21, 24, 25
	0.04	904.2	19, 21, 23, 24, 25	1,011	19, 21, 23, 24, 25	1,101.5	19, 21, 23, 24, 25	1,173.9	19, 21, 24, 25
	0.01	909.5	19, 21, 23, 24, 25	1,014.5	19, 21, 23, 24, 25	1,103.5	19, 21, 23, 24, 25	1,174.6	19, 21, 24, 25
	0	911.2	19, 21, 23, 24, 25	1,015.8	19, 21, 23, 24, 25	1,104.2	19, 21, 23, 24, 25	1,174.8	19, 21, 24, 25

Table 4 Multiple allocation results for AP dataset instances with 20 and 25 nodes

n	θ	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
		$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*
20	1	118.2	1, 8, 10, 16, 19, 20	146.3	1, 8, 10, 16, 19, 20	172.8	1, 8, 10, 11, 17, 20	195.6	1, 7, 8, 16, 17, 20
	0.1	120.7	1, 8, 10, 16, 19, 20	148.5	1, 8, 10, 11, 17, 20	174	1, 8, 10, 11, 17, 20	196.1	7, 8, 16, 17, 20
	0.07	121.6	1, 8, 10, 16, 19, 20	149.3	1, 8, 10, 11, 17, 20	174.4	1, 8, 10, 11, 17, 20	196.2	7, 8, 16, 17, 20
	0.04	122.6	1, 8, 10, 16, 19, 20	150	1, 8, 10, 16, 17, 20	174.9	1, 8, 10, 11, 17, 20	196.3	7, 8, 16, 17, 20
	0.01	123.7	1, 8, 10, 16, 19, 20	150.8	1, 8, 10, 16, 17, 20	175.4	1, 8, 10, 11, 17, 20	196.4	7, 8, 16, 17, 20
	0	124	1, 8, 10, 16, 19, 20	151.1	1, 8, 10, 16, 17, 20	175.6	1, 8, 10, 11, 17, 20	196.5	7, 8, 16, 17, 20
25	1	173.2	1, 5, 10, 16, 19, 21, 24	206.6	1, 5, 10, 16, 21, 24	233.6	1, 5, 10, 14, 21	254	1, 10, 14, 21
	0.1	175.5	1, 5, 10, 16, 19, 21, 24	208.1	1, 5, 10, 16, 21, 24	234.3	1, 5, 10, 14, 21	254.4	1, 10, 14, 21

Table 4 Multiple allocation results for AP dataset instances with 20 and 25 nodes (continued)

n	θ	$\alpha = 0.2$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.8$	
		$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*	$\bar{F}$	*Hub location*
25	0.07	176.2	1, 5, 10, 16, 19, 21, 24	208.6	1, 5, 10, 16, 21, 24	234.5	1, 5, 10, 14, 21	254.5	1, 10, 14, 21
	0.04	177.1	1, 5, 10, 16, 19, 21, 24	209.4	1, 5, 10, 16, 21, 24	234.8	1, 5, 10, 14, 21	254.6	1, 10, 14, 21
	0.01	178.2	1, 5, 10, 16, 19, 21, 24	209.9	1, 5, 10, 14, 21	235.2	1, 5, 10, 14, 21	254.7	1, 10, 14, 21
	0	178.5	1, 5, 10, 16, 19, 21, 24	210.1	1, 5, 10, 14, 21	235.3	1, 5, 10, 14, 21	254.8	1, 10, 14, 21

Tables 1 and 3 suggest that, given the value of α, the optimal solutions with symmetric OD flows ($\lambda = 1$) are always the same under different θ values for both single and multiple allocation cases. These results mean that the BRFCs are redundant for achieving balanced flows under symmetric OD flows. For any configuration of HSN, balanced round-trip flows on hub links can be achieved, since the routed flows are always same in round-trip for each link (hub link and spoke link) under symmetric OD flows. By contrast, to achieve balanced flows under asymmetric OD flows, the BRFCs are required to work. In fact, the asymmetry of OD flows causes an imbalance in round-trip flows, which leads to low utilisation of transportation resources. However, OD flows are usually asymmetric in reality due to the differences in regional economies and cultures. Therefore, HSN operators should focus more on balanced round-trip flows in the hub link transportation for improving resource utilisation.

For the single allocation case, we observe from Tables 1–2 and Figure 4 that the number of located hubs tends to decrease with a decrease in the value of θ under asymmetric OD flows ($\lambda > 1$). In particular, the number of located hubs may be reduced to one when θ approaches or equals zero. For example, for the AP instance with 25 nodes, the number of located hubs decreases from 7 to 1 when the value of θ decreases from 1 to 0 and the economy of scale factor is set to 0.4 (Figure 4). For the CAB instances, only node 20 is selected as a hub in Table 1 when $\theta = 0$ and $\lambda \in \{5, 7, 10\}$. The reason is that a lower number of located hubs can reduce the risk of not satisfying the BRFCs, and increase the number of spokes that can be flexibly adjusted and allocated. As a result, the solution of satisfying a smaller θ value is easier to achieve by optimising the allocation of spokes to hubs flexibly. In fact, for hub location problems with balanced flows, the solution with only one established hub is a feasible solution for the case of any θ ($\theta \geq 0$) value and asymmetric OD flows. This is because the OD flows are not required to be transferred between hubs. However, the case of one hub is the worse solution because its operating cost is the highest. Therefore, balanced round-trip flows on hub links can always be achieved for both single and multiple allocation cases.

However, for the multiple allocation case, the value of θ has a small impact on the number of located hubs under asymmetric OD flows (Tables 3 and 4). When θ is reduced, decreasing the number of located hubs can still be observed in some instances. But this trend is not obvious. For example, for the AP instance with 25 nodes, only one hub is reduced when the value of θ decreases from 1 to 0 and the economy of scale factor is set to 0.4 (Figure 5). For most instances, the hub location is same across different θ values when the values of α and λ are fixed. This is because the BRFCs with different θ values can be satisfied by flexibly allocating a spoke to different hubs and optimising spoke link flows, despite having the same hub location.

Figure 4 Number of located hubs for single allocation based on AP dataset instances with 25 nodes (see online version for colours)

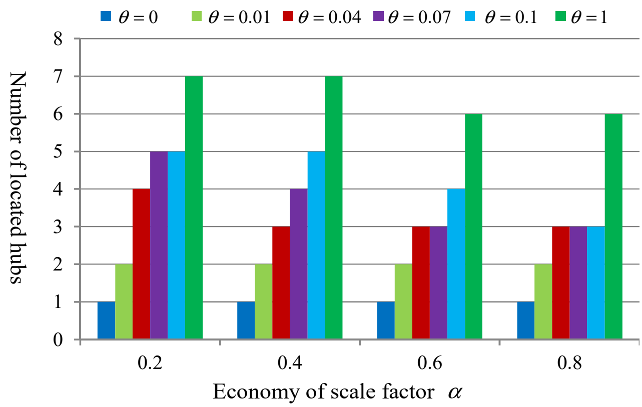
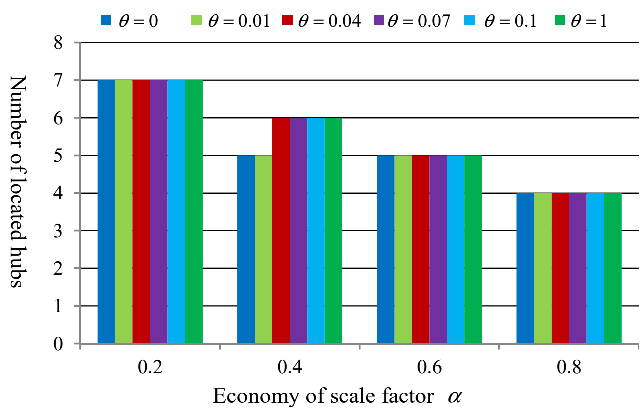


Figure 5 Number of located hubs for multiple allocation based on AP dataset instances with 25 nodes (see online version for colours)



Tables 1–4 suggest that the parameter of θ has a great impact on the operating cost of unit flow under asymmetric OD flows. When we just vary the parameter of θ, it is found that, for both single and multiple allocation cases, the operating cost of unit flow obtained based on the value of θ^* is greater than or equal to the operating cost of unit flow obtained based on the value of θ, if $\theta^* \leq \theta$ (Tables 1–4). The reason is that the optimal solution obtained based on the value of θ^* is a feasible solution for the case of the value of θ when $\theta^* \leq \theta$. Moreover, Figures 6 and 7 show the variation trend of the operating cost including the fixed setup cost and the transportation cost with the change of θ. When

θ is greater than or equal to a certain value (e.g., $\theta = 0.32563$ in Figure 6, and $\theta = 0.40502$ in Figure 7), the operating cost, the fixed setup cost and the transportation cost do not vary with the value of θ. This confirms Remark 1. While θ is smaller than a certain value (e.g., $\theta = 0.3256$ in Figure 6, and $\theta = 0.40502$ in Figure 7), as the decrease of θ value, the operating cost tends to increase because of the increase in the transportation cost, and the fixed setup cost tends to decrease due to the fewer hubs. These results show that increasing the traditional operating cost may achieve balanced round-trip flows on hub links. In addition, we note that the operating cost of the single allocation case is higher than that of the multiple allocation case, since the single allocation solution is a feasible solution for the multiple allocation case.

Figure 6 Variation trend in costs for single allocation (AP dataset with 25 nodes, $\alpha = 0.4$) (see online version for colours)

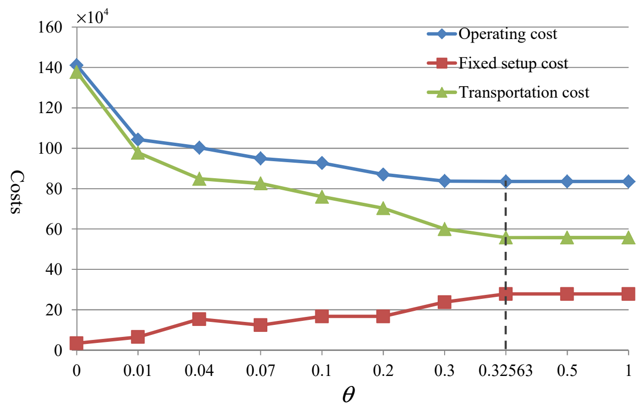
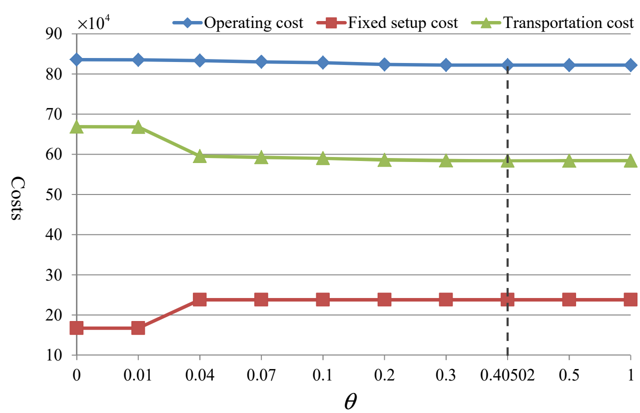


Figure 7 Variation trend in costs for multiple allocation (AP dataset with 25 nodes, $\alpha = 0.4$) (see online version for colours)



For the impacts of balanced flows on HSN structures, we observe that the optimal HSN structure tends to differ under different values of θ (Tables 1–4 and Figures 8–9). For the single allocation case, Figures 8(a) and 8(b) show a difference between unbalanced and balanced flows in hub locations and allocation of spokes to hubs. There is obviously regional or geographical advantage for the allocation of spokes to hubs under unbalanced flows. For example, Figure 8(a) depicts that all spokes are allocated to the relatively near hubs, and hubs 19 and 25 are responsible for the western and northeastern parts of the USA, respectively. Node 21 seems to be the hub for the central part of the USA, whereas hub 24 is responsible for the southeastern part of the country.

However, for balanced flows, there is basically no geographical advantage for the allocation of spokes to hubs. For example, spokes 7 and 23 are allocated to the relatively distant hubs 20 and 21 in Figure 8b, respectively. This phenomenon can be attributed to the fact that the BRFCs make the configuration of HSN different from the case of unbalanced flows. As a result, some spokes are allocated to the distant hubs to satisfy the balancing requirements, which leads to an increase in the routing cost for OD flows.

Moreover, the optimal hub locations and allocation of spokes to hubs are almost different with changes in the value of θ [Figures 8(c) and 8(d)]. For example, compared to the HSN returned by solving the instance with θ = 0.07, α = 0.4, and λ = 5 [Figure 8(c)], the allocation of node 5 to hubs is changed in the HSN returned by solving the instance with θ = 0.04, α = 0.4, and λ = 5 [Figure 8(d)], despite having the same hub nodes 19, 21, and 25.

Figure 8 HSNs for single allocation based on the CAB dataset (α = 0.4 and λ = 5), (a) HSN with unbalanced flows (θ = 1) (b) HSN with balanced flows (θ = 0.1) (c) HSN with balanced flows (θ = 0.07) (d) HSN with balanced flows (θ = 0.04) (see online version for colours)

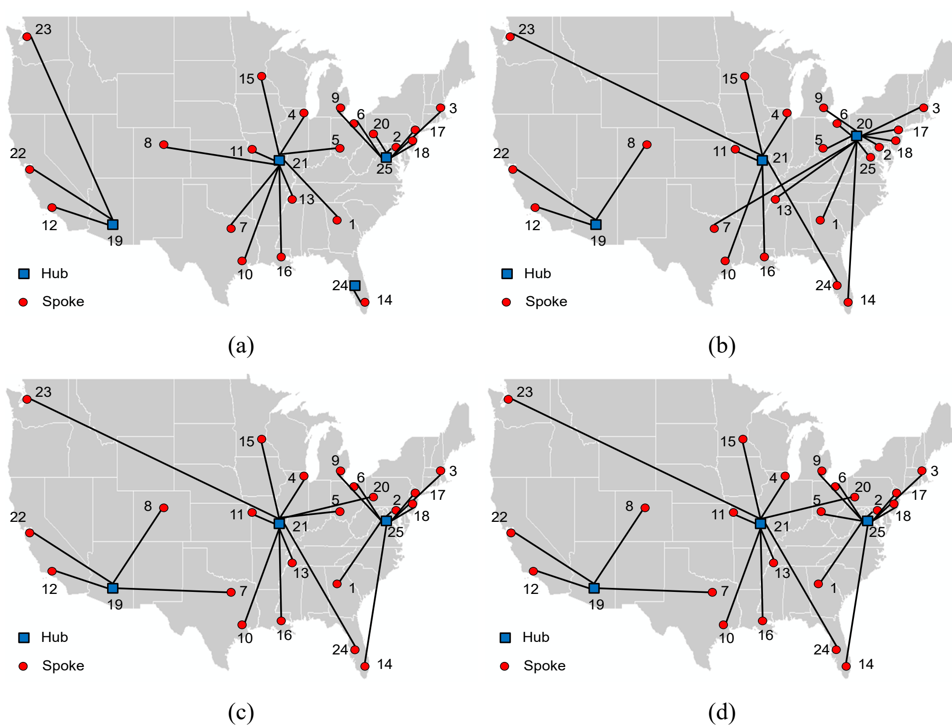


Figure 9 HSNs for multiple allocation based on the CAB dataset ($\alpha = 0.4$ and $\lambda = 5$), (a) HSN with unbalanced flows ($\theta = 1$) (b) HSN with balanced flows ($\theta = 0.1$) (c) HSN with balanced flows ($\theta = 0.04$) (d) HSN with balanced flows ($\theta = 0$) (see online version for colours)

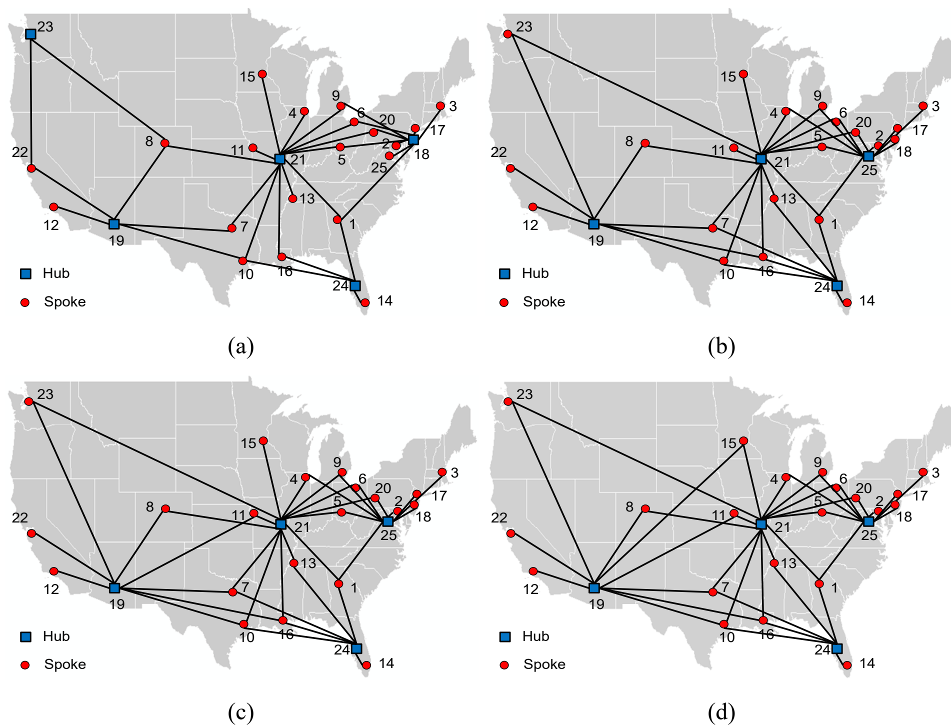
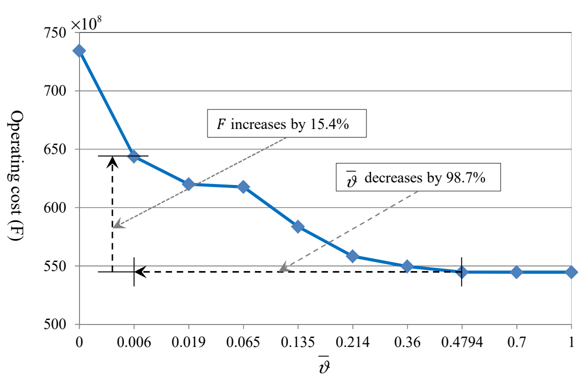


Figure 10 Trade-off between the operating cost and the entire unbalanced round-trip degree ($\bar{\vartheta}$) of the hub links for single allocation (CAB dataset with 25 nodes, $\alpha = 0.6$, and $\lambda = 10$) (see online version for colours)



For the multiple allocation case, Figures 9(a) and 9(b) show a difference between unbalanced and balanced flows in hub locations and allocation of spokes to hubs. There is regional or geographical advantage for the allocation of spokes to hubs under unbalanced and balanced flows. Spokes are in general allocated to the hubs that are relatively near. Meanwhile, as the value of θ decreases, the allocation of spokes to hubs and optimal spoke link flows tend to differ. For example, when θ reduces from 0.04 to 0, the allocation of spoke 15 changes from hub 21 to hubs 19 and 21, despite having the same hub nodes 19, 21, 24, and 25 [Figures 9(c) and 9(d)]. These results demonstrate that balanced flows can be achieved by optimising the HSN structure and the routed flows.

Figure 11 Trade-off between the operating cost and the entire unbalanced round-trip degree ($\overline{\vartheta}$) of the hub links for multiple allocation (CAB dataset with 25 nodes, $\alpha = 0.6$, and $\lambda = 10$) (see online version for colours)

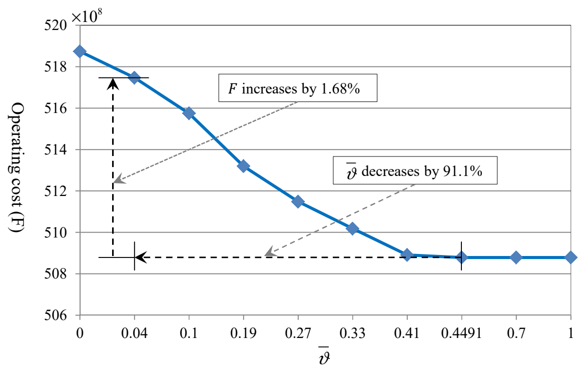
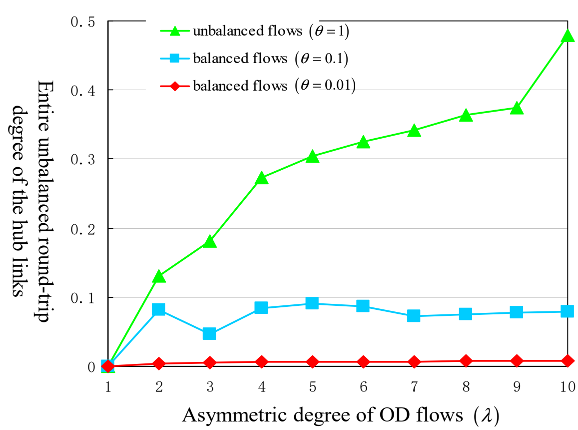
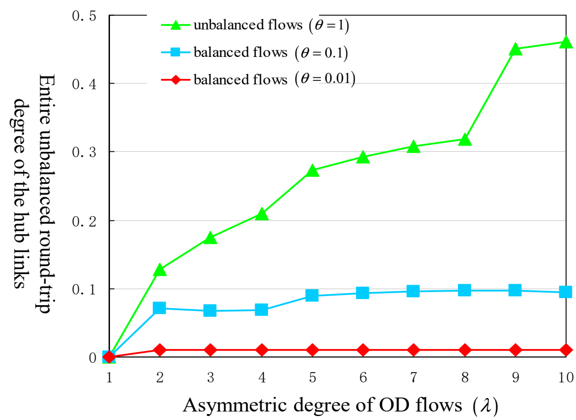


Figure 12 Entire unbalanced round-trip degree ($\overline{\vartheta}$) of the hub links for single allocation under different λ and θ values (see online version for colours)



Furthermore, the configuration of the HSN is also changed to accommodate the higher degree of asymmetry in OD flows. For example, Table 1 shows that hub locations change from nodes 13, 19 and 20 to nodes 19, 21, and 25 when the asymmetric degree of OD flows increases from 2 to 10 and $\theta = 0.04$, $\alpha = 0.4$. Even though the optimal solutions tend to differ under different parameters, some hub nodes are common across the results (Tables 1–4). For example, Tables 1 and 3 show that node 19 is almost always chosen as a hub under different values of θ, α and λ. Similarly, either node 20 or node 21 is usually present in the hub set in Table 1 and either node 21 or node 24 is usually present in the hub set in Table 3. These results suggest that some hub locations are robust to different balancing requirements, the economy of scale factors and also to the asymmetric OD flows. This is primarily due to the geographical locations and the setup cost size of hubs. As we noted, in the CAB dataset, node 19 generates the second lowest and the third lowest setup cost for establishing a hub when $\lambda \in \{1, 2, 3\}$ and $\lambda \in \{5, 7, 10\}$, respectively. Therefore, identifying hub locations which are robust to different parameters can better provide a basis for operators in designing an HSN with balanced flows.

Figure 13 Entire unbalanced round-trip degree ($\bar{\vartheta}$) of the hub links for multiple allocation under different λ and θ values (see online version for colours)



Additionally, to further analyse the impacts of balanced and unbalanced flows on optimal solutions, the entire unbalanced round-trip degree ($\bar{\vartheta}$) of the hub links is defined by equation (27), where p represents the number of located hubs.

$$\bar{\vartheta} = \frac{\sum_{k,l,k<l} \vartheta_{k,l}}{0.5 \cdot p \cdot (p-1)} \tag{27}$$

The trade-off curves between the operating cost and the entire unbalanced round-trip degree of the hub links are plotted in Figures 10 and 11. Figures 10 and 11 indicate a negative relationship between the operating cost and the entire unbalanced round-trip degree of the hub links. This means that increasing the traditional operating cost decreases the entire unbalanced round-trip degree of the hub links. For the single allocation case (Figure 10), the entire unbalanced round-trip degree of the hub links can be significantly decreased with a small increase in the operating cost. For example, the entire unbalanced round-trip degree of the hub links can be reduced from 0.4794 to 0.006, or 98.7%, with a small increase in the operating cost, from 544.635×10^8 to 643.605×10^8 (15.4%). In particular, when the entire unbalanced round-trip degree of the hub links is reduced from 1 to 0.4794 (Figure 10), the operating cost does not increase because the BRFCs are not binding. Such a phenomenon can also be observed in the multiple allocation case (Figure 11). For example, the entire unbalanced round-trip degree of the hub links can be reduced from 0.4491 to 0.04, or 91.1%, with a very small increase in the operating cost, from 508.787×10^8 to 517.467×10^8 (1.68%). Based on managerial preferences for the operating cost and the entire unbalanced round-trip degree of the hub links, the decision-maker can achieve a compromise solution by adjusting the θ value. If decision-makers prefer to minimise the operating cost, then a larger θ value should be set. Conversely, if decision-makers prefer to minimise the entire unbalanced round-trip degree of the hub links, then a smaller θ value should be set. However, to effectively determine the θ value, HSN operators should weigh the costs of balanced round-trip flows against the benefits it will bring.

Figures 12 and 13 show that the entire unbalanced round-trip degree of the hub links tends to increase with increase in the asymmetric degree of OD flows under unbalanced flows. Meanwhile, the entire unbalanced round-trip degree of the hub links is equal to zero for both balanced and unbalanced flows due to the symmetry of OD flows. This confirms that the imbalance of regional OD flows is the main cause of unbalanced round-trip flows. However, with balanced flows, the entire unbalanced round-trip degree of the hub links can be always limited in a given level to decrease the empty-trip rate of vehicles and increase the full-load rate of vehicles, due to the BRFCs.

3.3 Analysing aggregated hub link flows

Campbell (2013) examined the optimal solutions of hub location models and showed that the flows on spoke links frequently exceed the flows between hubs, which violates the premise underlying the model for economies of scale. The poor modelling of economies of scale appears to be quite common across basic hub location problems, especially with real-world datasets. To analyse the impacts of balanced flows on the aggregated hub link flows and the presumed economies of scale, we report the average percentage of the spokes with a flow greater than the hub link flow (Ω), the percentage of the hub links with a flow less than the maximum flow on any spoke (Ψ), and the average flows on the hub links (Φ) for both single and multiple allocation cases in Tables 5–8, as highlighted by Campbell (2013). Ω and Ψ are two measures of the degree to which the flows on spokes exceed the flows on hub links. To determine the value of Ω, for each hub link we calculate the percentage of the spokes with a flow exceeding that on the hub link, and then average these across the $p(p-1)$ hub links. Note that only when the values of Ω and Ψ are 0%, all hub link flows are greater than all spoke flows.

For both single and multiple allocation cases (Tables 5 and 7), there are no differences between unbalanced and balanced flows in the values of Ω, Ψ and Φ, under the symmetric OD flows. For the single allocation case (Tables 5 and 6), it can be observed that the value of Ω returned by considering balanced flows is significantly smaller than that of unbalanced flows, and the value of Φ returned by considering balanced flows is significantly greater than that of unbalanced flows under the asymmetric OD flows, except for the case of $\theta = 0$. As the value of θ decreases, the value of Ω tends to decrease and the value of Φ tends to increase in given the values of λ and α. For example, for the AP instance with 20 nodes, Ω decreases from 40% to 0% and Φ increases from 0.9663 to 8.6989 when the value of θ decreases from 1 to 0.01 and the economy of scale factor is set to 0.4 (Table 6). This is attributed to the fact that the BRFCs tend to spread the flows more evenly over the hub links, which may help to reduce the frequency of small hub link flows. As a result, the average percentage of the spokes with a flow greater than the hub link flow is decreased. The results for the single allocation case with balanced flows (except for the case of one established hub) show a better modelling of economies of scale as compared to the case of unbalanced flows.

Table 5 Comparisons of hub link flows and spoke flows for single allocation (CAB dataset, 25 nodes)

λ	θ	$\alpha = 0.2$			$\alpha = 0.4$			$\alpha = 0.6$			$\alpha = 0.8$		
		Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)
1	1	57.50	100	2.9256	34.92	100	4.4586	15.15	100	7.7733	15.15	100	7.7733
	0.1	57.50	100	2.9256	34.92	100	4.4586	15.15	100	7.7733	15.15	100	7.7733
	0.07	57.50	100	2.9256	34.92	100	4.4586	15.15	100	7.7733	15.15	100	7.7733
	0.04	57.50	100	2.9256	34.92	100	4.4586	15.15	100	7.7733	15.15	100	7.7733
	0.01	57.50	100	2.9256	34.92	100	4.4586	15.15	100	7.7733	15.15	100	7.7733
	0	57.50	100	2.9256	34.92	100	4.4586	15.15	100	7.7733	15.15	100	7.7733
2	1	56.50	100	4.3884	56.50	100	4.3884	61.13	100	4.3403	26.14	100	8.3160
	0.1	46.23	100	5.9751	48.41	100	5.7449	17.05	100	11.333	21.97	100	8.5491
	0.07	15.91	100	9.9676	20.45	100	9.1145	20.45	100	9.1145	20.45	100	9.1145
	0.04	12.50	100	11.965	13.64	100	11.474	3.26	100	18.129	3.26	100	18.129
	0.01	12.88	100	11.782	12.88	100	11.782	4.35	100	17.282	4.35	100	17.282
	0	--	--	--	--	--	--	--	--	--	--	--	--
3	1	38.10	100	8.8337	50.60	100	7.4068	50.60	100	7.4068	50.00	100	7.4778
	0.1	34.85	100	12.688	18.94	100	12.153	18.94	100	12.153	18.94	100	12.153
	0.07	16.67	100	15.036	16.67	100	15.036	5.43	100	22.943	5.43	100	22.943
	0.04	11.36	100	15.954	17.05	100	14.173	19.70	100	12.917	20.08	100	12.424
	0.01	12.12	100	17.253	4.35	100	23.043	4.35	100	23.043	4.35	100	23.043
	0	--	--	--	--	--	--	--	--	--	--	--	--

Note: That '--' represents the null value because there are no hub links.

Table 5 Comparisons of hub link flows and spoke flows for single allocation (CAB dataset, 25 nodes) (continued)

λ	θ	$\alpha = 0.2$			$\alpha = 0.4$			$\alpha = 0.6$			$\alpha = 0.8$		
		Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)
5	1	62.50	100	8.1895	42.86	100	12.751	44.25	100	12.637	44.25	100	12.637
	0.1	25.38	100	19.877	15.91	100	21.932	15.91	100	21.932	15.91	100	21.932
	0.07	16.29	100	22.949	10.61	100	23.931	4.35	100	36.257	4.35	100	36.257
	0.04	12.50	100	23.630	10.98	100	23.630	10.98	100	23.630	2.17	100	49.162
	0.01	12.50	100	27.063	4.35	100	34.564	4.35	100	34.564	4.35	100	34.564
	0	--	--	--	--	--	--	--	--	--	--	--	--
7	1	62.00	100	10.920	66.13	100	10.488	66.13	100	10.488	44.84	100	16.849
	0.1	35.52	100	18.538	10.61	100	31.909	12.50	100	31.764	20.83	100	25.742
	0.07	10.61	100	31.909	10.61	100	31.909	12.12	100	31.383	2.17	100	63.092
	0.04	10.98	100	31.507	10.98	100	31.507	10.98	100	31.507	2.17	100	65.550
	0.01	12.12	100	32.386	12.12	100	32.386	4.35	100	46.086	2.17	100	65.831
	0	--	--	--	--	--	--	--	--	--	--	--	--
10	1	62.00	100	15.015	65.63	100	14.421	65.63	100	14.421	65.63	100	14.421
	0.1	10.61	100	43.876	10.61	100	43.876	11.74	100	43.676	21.59	100	35.395
	0.07	10.61	100	43.876	10.61	100	43.876	10.61	100	43.876	4.35	100	66.472
	0.04	10.98	100	43.323	10.98	100	43.323	10.98	100	43.323	10.98	100	43.323
	0.01	11.74	100	44.531	11.74	100	44.531	11.74	100	44.531	2.17	100	99.719
	0	--	--	--	--	--	--	--	--	--	--	--	--

Note: That '--' represents the null value because there are no hub links.

Table 6 Comparisons of hub link flows and spoke flows for single allocation (AP dataset, 20 and 25 nodes)

n	*θ*	*α* = 0.2			*α* = 0.4			*α* = 0.6			*α* = 0.8		
		Ω (%)	*Ψ* (%)	*Φ*	*Ω* (%)	*Ψ* (%)	*Φ*	*Ω* (%)	*Ψ* (%)	*Φ*	*Ω* (%)	*Ψ* (%)	*Φ*
20	1	37.38	100	0.9678	40.00	100	0.9663	47.26	100	0.9656	47.26	100	0.9656
	0.1	36.00	100	1.2874	39.17	100	1.2874	40.00	100	1.2844	40.00	100	1.2844
	0.07	16.67	100	3.6159	25.49	100	3.4169	25.49	100	3.4169	18.14	100	3.3358
	0.04	23.04	100	3.3358	18.63	100	3.3326	18.14	100	3.3358	18.14	100	3.3358
	0.01	0	0	8.6989	0	0	8.6989	0	0	8.9276	--	--	--
	0	--	--	--	--	--	--	--	--	--	--	--	--
25	1	48.41	100	0.8927	53.44	100	0.8927	53.60	100	1.1174	53.60	100	1.1174
	0.1	45.00	100	1.6330	45.00	100	1.6330	39.09	100	2.1229	20.45	100	3.2904
	0.07	44.25	100	1.6816	37.30	100	2.2244	21.97	100	3.2866	21.97	100	3.2866
	0.04	36.11	100	2.6318	14.77	100	4.5423	14.02	100	4.5456	20.45	100	3.3115
	0.01	2.17	100	7.2804	2.17	100	7.2804	2.17	100	7.2804	2.17	100	7.2804
	0	--	--	--	--	--	--	--	--	--	--	--	--

Note: That '--' represents the null value because there are no hub links.

Table 7 Comparisons of hub link flows and spoke flows for multiple allocation (CAB dataset, 25 nodes)

λ	θ	$\alpha = 0.2$			$\alpha = 0.4$			$\alpha = 0.6$			$\alpha = 0.8$		
		Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)
1	1	53.75	100	2.7132	23.04	100	2.9576	18.92	100	3.3078	28.03	100	2.1689
	0.1	53.75	100	2.7132	23.04	100	2.9576	18.92	100	3.3078	28.03	100	2.1689
	0.07	53.75	100	2.7132	23.04	100	2.9576	18.92	100	3.3078	28.03	100	2.1689
	0.04	53.75	100	2.7132	23.04	100	2.9576	18.92	100	3.3078	28.03	100	2.1689
	0.01	53.75	100	2.7132	23.04	100	2.9576	18.92	100	3.3078	28.03	100	2.1689
	0	53.75	100	2.7132	23.04	100	2.9576	18.92	100	3.3078	28.03	100	2.1689
2	1	53.02	100	4.0698	39.78	100	4.1433	16.89	100	4.9616	28.03	100	3.2533
	0.1	36.46	100	5.4507	38.62	100	4.1474	16.89	100	5.0037	25.93	100	3.3790
	0.07	35.23	100	5.4877	38.44	100	4.1437	18.04	100	4.9789	25.84	100	3.3812
	0.04	35.61	100	5.5088	37.37	100	4.1322	17.56	100	4.9921	25.84	100	3.3812
	0.01	36.97	100	5.5253	38.28	100	4.0972	16.44	100	5.0497	26.09	100	3.3317
	0	36.97	100	5.5345	38.28	100	4.0943	15.98	100	5.0685	26.09	100	3.3314
3	1	36.57	100	7.4472	33.20	100	6.5034	36.51	100	4.6746	31.55	100	4.4863
	0.1	35.12	100	7.2544	32.35	100	6.2387	36.29	100	4.7722	29.96	100	4.3347
	0.07	34.65	100	7.1729	32.96	100	6.2065	37.00	100	4.7719	30.49	100	4.3240
	0.04	34.36	100	7.2548	33.45	100	6.1766	37.39	100	4.7364	30.84	100	4.2932
	0.01	34.97	100	7.2731	34.10	100	6.2198	37.34	100	4.6668	30.71	100	4.2289
	0	35.12	100	7.2582	33.59	100	6.0828	36.80	100	4.6613	30.40	100	4.1866

Table 7 Comparisons of hub link flows and spoke flows for multiple allocation (CAB dataset, 25 nodes) (continued)

λ	θ	$\alpha = 0.2$			$\alpha = 0.4$			$\alpha = 0.6$			$\alpha = 0.8$		
		Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)	Ω (%)	Ψ (%)	Φ ($\times 10^5$)
5	1	59.07	100	6.8355	53.26	100	5.9857	31.94	100	7.6178	29.61	100	6.3473
	0.1	55.45	100	6.7361	34.47	100	9.0282	30.35	100	7.3140	29.34	100	5.9945
	0.07	56.18	100	6.8170	33.83	100	8.9785	30.18	100	7.3025	28.87	100	5.9137
	0.04	55.54	100	6.8045	34.58	100	8.8692	31.04	100	7.1636	28.75	100	5.8506
	0.01	53.88	100	6.7630	34.56	100	8.6744	31.46	100	7.1235	33.49	100	4.8720
	0	34.80	100	10.797	34.56	100	8.6292	31.67	100	7.1028	33.49	100	4.8871
7	1	57.77	100	8.7820	54.53	100	8.1679	51.34	100	6.2806	28.03	100	8.4631
	0.1	51.84	100	9.3024	52.36	100	7.0967	30.97	100	9.6519	28.58	100	7.8719
	0.07	56.02	100	9.3472	51.15	100	7.0097	31.30	100	9.5814	29.21	100	7.7610
	0.04	57.31	100	9.4089	52.29	100	6.9716	31.69	100	9.4029	28.43	100	7.6899
	0.01	51.29	100	9.1803	49.68	100	6.7981	32.91	100	9.4141	28.71	100	7.7320
	0	56.90	100	8.2304	34.30	100	11.339	32.91	100	9.4684	29.54	100	7.4240
10	1	57.77	100	12.075	54.06	100	11.231	50.98	100	8.6358	47.50	100	7.1788
	0.1	53.21	100	12.812	52.22	100	9.5736	49.83	100	7.9156	29.33	100	10.645
	0.07	54.91	100	12.826	51.16	100	9.5498	48.30	100	7.8071	28.04	100	10.543
	0.04	54.74	100	12.734	50.80	100	9.3168	50.00	100	7.7914	28.24	100	10.485
	0.01	52.83	100	12.357	52.33	100	9.0944	50.06	100	7.7809	26.76	100	10.444
	0	54.26	100	11.072	52.80	100	9.0531	49.67	100	7.7621	30.17	100	10.023

Table 8 Comparisons of hub link flows and spoke flows for multiple allocation (AP dataset, 20 and 25 nodes)

n	*θ*	*α* = 0.2			*α* = 0.4			*α* = 0.6			*α* = 0.8		
		Ω (%)	*Ψ* (%)	*Φ*	*Ω* (%)	*Ψ* (%)	*Φ*	*Ω* (%)	*Ψ* (%)	*Φ*	*Ω* (%)	*Ψ* (%)	*Φ*
20	1	30.56	100	0.8926	25.68	100	0.8735	27.15	100	0.8624	24.02	100	0.6019
	0.1	25.16	100	0.9058	28.13	100	0.9035	26.60	100	0.8576	25.76	100	0.7851
	0.07	25.16	100	0.9036	28.33	100	0.8944	26.24	100	0.8539	25.49	100	0.7791
	0.04	26.10	100	0.9003	25.80	100	0.8864	26.85	100	0.8369	24.17	100	0.7802
	0.01	25.50	100	0.8985	25.80	100	0.8827	26.97	100	0.8163	23.82	100	0.7799
	0	26.05	100	0.8986	26.23	100	0.8802	27.90	100	0.8138	23.84	100	0.7747
25	1	38.50	100	0.8345	36.49	100	0.9924	34.84	100	1.2255	33.43	100	1.1258
	0.1	34.92	100	0.8418	34.04	100	0.9803	33.51	100	1.1957	30.59	100	1.1750
	0.07	33.74	100	0.8357	32.28	100	0.9824	33.16	100	1.1870	30.40	100	1.1682
	0.04	35.21	100	0.8368	30.61	100	0.9912	33.31	100	1.1788	30.71	100	1.0874
	0.01	35.40	100	0.8385	35.08	100	1.3325	33.38	100	1.1683	31.34	100	1.0816
	0	33.94	100	0.8358	35.17	100	1.3321	33.62	100	1.1644	31.25	100	1.0801

For the multiple allocation case (Tables 7 and 8), although the value of Φ has a small reduction under balanced flows in most instances, the value of Ω returned by considering balanced flows is still smaller than that of unbalanced flows, which accounts for 82.14% of all instances with balanced flows. Only 17.86% of all instances with balanced flows have a higher Ω value, but these instances do not significantly affect the modelling of economies of scale, since the increment of Ω value is very small. Moreover, the number of spoke links obtained from the instances with balanced flows is greater than that of unbalanced flows in general, which decreases and splits the larger spoke flows to some degree. Therefore, the low frequency of small hub link flows along with the reduced spoke flows reduces the value of Ω. A better fit to the modelling of economies of scale can be achieved in the multiple allocation case with balanced flows.

However, since nearby nodes often exchange large flows, there are still a few spokes with large flows (see the Ψ value of Tables 5–8). For both single and multiple allocation cases, the value of Ω tends to increase with the increase in the number of located hubs (Tables 1–8), a result that is consistent with Campbell (2013). However, under the consideration of balanced flows, fewer hubs are located in general. Therefore, from the perspective of better modelling of economies of scale, it makes sense to consider the balanced flows in the design of HSN. However, the trade-off between the benefits yielded from the better modelling of economies of scale and the costs generated from the achieving of balanced flows requires further study.

3.4 Performance of CPLEX

To analyse the performance of CPLEX under balanced flows, experiments with different class sizes (n = 15, 20, 25) are conducted. Tables 9 and 10 report the runtimes of CPLEX 12.6 for solving CAB and AP dataset instances, respectively. All instances are solved with a 0% optimality gap stopping rule and the α value is set to 0.6.

For both single and multiple allocation cases, we observe from Tables 9 and 10 that the solution times for CPLEX increase with the instance class size. In particular, with respect to the single allocation case and asymmetric OD flows, the solution times of CPLEX markedly increase with the increase of the instance class size under balanced flows. For example, for the instance of $\lambda = 7$ and $\theta = 0$, the solution time increases from 2.51 to 14,761.55 seconds when the number of nodes for CAB dataset instances increases from 15 to 25 (Table 9). With symmetric OD flows, CPLEX returns the same solution time under different θ values because the BRFCs are not binding. For example, for CAB dataset instances (n = 25), the solution time is 3.98 seconds across all different θ values with multiple allocation case. Under asymmetric OD flows, by comparing the solution times of unbalanced flows ($\theta = 1$) and balanced flows ($0 \leq \theta < 1$), it is observed that the solution time of balanced flows is always larger than or equal to that of unbalanced flows for both single and multiple allocation instances. Meanwhile, as the value of θ decreases, the solution time tends to increase for single allocation instances under asymmetric OD flows. Especially, this trend is particularly evident when the instance class size is large. For example, for AP dataset instances (n = 25), the solution time increases from 1.72 to 16,758.91 seconds when the θ value decreases from 1 to 0 (Table 10). With the increase of the asymmetric degree of OD flows, the solution time tends to increase for single allocation instances in given the value of θ. For example, for CAB dataset instances (n = 20), the solution time increases from 7.55 to 214.36 seconds when the λ value increases from 3 to 10 and the θ value is set to 0.1 (Table 9). However, such phenomena are not observed in the multiple allocation case. The change of θ ($0 \leq \theta < 1$) value has a small impact on the variation of solution times of multiple allocation instances with asymmetric OD flows. For example, when the value of θ varies within the range of 0 to 0.1, the solution time only fluctuates within the range of 7.86 to 15.16 seconds with AP dataset instances (n = 25). Moreover, the increase of the asymmetric degree of OD flows does not result in the increase of the solution time for multiple allocation instances in given the value of θ.

Table 9 Runtimes of CPLEX for solving CAB dataset instances (sec.)

λ	θ	*Single allocation*			*Multiple allocation*		
		n = 15	*n = 20*	*n = 25*	*n = 15*	*n = 20*	*n = 25*
1	1	0.51	1.48	4.46	0.25	1.31	3.98
	0.1	0.51	1.48	4.46	0.25	1.31	3.98
	0.04	0.51	1.48	4.46	0.25	1.31	3.98
	0.01	0.51	1.48	4.46	0.25	1.31	3.98
	0	0.51	1.48	4.46	0.25	1.31	3.98
3	1	0.54	1.50	7.89	0.20	0.89	10.05
	0.1	2.03	7.55	910.19	0.69	2.92	72.40
	0.04	1.06	67.91	8,612.15	0.69	3.40	57.63
	0.01	1.01	35.57	9,734.18	0.67	2.54	26.32
	0	0.81	1,628.43	12,340.26	0.67	2.40	17.61
7	1	0.62	1.53	14.93	0.22	0.56	3.35
	0.1	4.01	90.65	2,341.38	0.61	3.90	12.54
	0.04	1.87	46.58	9,102.38	0.56	2.67	13.71
	0.01	2.31	457.82	11,453.14	0.56	3.04	11.12
	0	2.51	7,836.63	14,761.55	0.48	2.65	7.69
10	1	0.73	1.89	4.87	0.22	0.61	2.20
	0.1	2.98	214.36	3,427.21	0.69	2.64	10.70
	0.04	1.45	490.25	11,273.56	0.51	2.29	10.47
	0.01	3.49	476.82	14,278.91	0.47	2.25	8.89
	0	1.44	8,320.73	18,456.37	0.51	2.12	6.83

Table 10 Runtimes of CPLEX for solving AP dataset instances (sec.)

θ	*Single allocation*		*Multiple allocation*	
	n = 20	*n = 25*	*n = 20*	*n = 25*
1	0.58	1.72	1.40	2.31
0.1	147.44	2,951.29	9.14	10.99
0.04	162.23	2,509.53	10.42	12.19
0.01	798.24	3,460.40	13.98	15.16
0	1,813.33	16,758.91	6.19	7.86

Additionally, to further show the performance of CPLEX for large-scale instances, we solve AP dataset instances (n = 40) under balanced flows. It is reported that CPLEX can not solve the instances with low θ values to 0% optimality because it ran out of memory. Therefore, to solve large-scale HSN design problems with balanced flows, developing fast and effective solution algorithms such as a Lagrangian heuristic algorithm for the formulated models is a necessity in future work.

4 Conclusions

This work studies an extension of classical single and multiple allocation hub location problems in which the balanced round-trip flows on hub links are considered in HSN design. The main goal of balanced round-trip flows on hub links is to decrease the empty-trip rate of vehicles and increase the full-load rate of vehicles. Mixed-integer programming models are established to cope with balanced flows for single and multiple allocation cases. To analyse the impacts of balanced flows on HSN configurations and aggregated hub link flows, the extensive computational analysis with more than 300 instances on the CAB and AP datasets is performed.

Experimental results show that the consideration of balanced flows in HSN design greatly affects the configuration of the HSN, including hub location and allocation of spokes to hubs. The number of located hubs for the single allocation case appears to be insensitive to changes in the economy of scale factor. However, as expected, increasing the economy of scale factor in general leads to fewer hubs for the multiple allocation case. Furthermore, for the single allocation case, the number of located hubs tends to decrease with a decrease in the value of θ under asymmetric OD flows. However, this trend is not obvious in the multiple allocation case. Although optimal hub locations tend to differ under different parameters, several common hubs are robust to different balancing requirements, economy of scale factors, and also to the asymmetric OD flows. Moreover, for both single and multiple allocation cases, the entire unbalanced round-trip degree of the hub links can be significantly decreased with a small increase in the traditional operating cost. This means that the utilisation rate of transportation resources can be significantly improved with a small increase in the traditional operating cost. In addition, the analysis also shows that an imbalance in regional OD flows causes an imbalance in round-trip flows. Under symmetric OD flows, balanced round-trip flows on hub links can always be achieved in any configuration of HSN. A better fit to the modelling of economies of scale can be achieved by considering balanced flows. For both single and multiple allocation cases, the solution times for CPLEX increase with the instance class size. With the decrease of the value of θ, the solution time tends to increase for single allocation instances under asymmetric OD flows.

This work represents a new research direction in handling hub location problems. It also represents a contribution to the better modelling of economies of scale.

There are several areas that future work can focus on. First, the present work can be extended to different variants of hub location problems such as incomplete hub network design, capacitated, hierarchical, and multimodal hub location problems. Second, in a dynamic and competitive business environment, demand uncertainty should be incorporated into hub location problems with balanced flows. Third, to further improve operational efficiency, balanced flows on spoke links should also be considered in the design of HSN. Finally, the trade-off between the benefits yielded by balanced flows and the costs generated by such operations may need to be further investigated.

References

Alkaabneh, F., Diabat, A. and Elhedhli, S. (2019) 'A Lagrangian heuristic and GRASP for the hub-and-spoke network system with economies-of-scale and congestion', *Transportation Research Part C: Emerging Technologies*, Vol. 102, pp.249–273.

Alumur, S. and Kara, B.Y. (2008) 'Network hub location problems: the state of the art', *European Journal of Operational Research*, Vol. 190, No. 1, pp.1–21.

Alumur, S.A., Kara, B.Y. and Karasan, O.E. (2009) 'The design of single allocation incomplete hub networks', *Transportation Research Part B: Methodological*, Vol. 43, No. 10, pp.936–951.

Alumur, S.A., Kara, B.Y. and Karasan, O.E. (2012) 'Multimodal hub location and hub network design', *Omega*, Vol. 40, No. 6, pp.927–939.

An, Y., Zhang, Y. and Zeng, B. (2015) 'The reliable hub-and-spoke design problem: models and algorithms', *Transportation Research Part B: Methodological*, Vol. 77, pp.103–122.

Aykin, T. (1995) 'Networking policies for hub-and-spoke systems with application to the air transportation system', *Transportation Science*, Vol. 29, No. 3, pp.201–221.

Azizi, N., Chauhan, S., Salhi, S. and Vidyarthi, N. (2016) 'The impact of hub failure in hub-and-spoke networks: mathematical formulations and solution techniques', *Computers & Operations Research*, Vol. 65, pp.174–188.

Azizi, N., Vidyarthi, N. and Chauhan, S.S. (2018) 'Modelling and analysis of hub-and-spoke networks under stochastic demand and congestion', *Annals of Operations Research*, Vol. 264, Nos. 1–2, pp.1–40.

Berman, O., Drezner, Z. and Wesolowsky, G.O. (2007) 'The transfer point location problem', *European Journal of Operational Research*, Vol. 179, No. 3, pp.978–989.

Campbell, J.F. (1994) 'Integer programming formulations of discrete hub location problems', *European Journal of Operational Research*, Vol. 72, No. 2, pp.387–405.

Campbell, J.F. (1996) 'Hub location and the p-hub median problem', *Operations Research*, Vol. 44, No. 6, pp.923–935.

Campbell, J.F. (2009) 'Hub location for time definite transportation', *Computers & Operations Research*, Vol. 36, No. 12, pp.3107–3116.

Campbell, J.F. (2013) 'Modeling economies of scale in transportation hub networks', *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS)*, IEEE, Wailea, Maui, HI, USA, pp.1154–1163.

Campbell, J.F. and O'Kelly, M.E. (2012) 'Twenty-five years of hub location research', *Transportation Science*, Vol. 46, No. 2, pp.153–169.

Cetiner, S., Sepil, C. and Süral, H. (2010) 'Hubbing and routing in postal delivery systems', *Annals of Operations Research*, Vol. 181, No. 1, pp.109–124.

Chen, H., Campbell, A.M. and Thomas, B.W. (2008) 'Network design for time-constrained delivery', *Naval Research Logistics*, Vol. 55, No. 6, pp.493–515.

Contreras, I., Cordeau, J-F. and Laporte, G. (2011) 'The dynamic uncapacitated hub location problem', *Transportation Science*, Vol. 45, No. 1, pp.18–32.

Correia, I., Nickel, S. and Saldanha-Da-Gama, F. (2010a) '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.

Correia, I., Nickel, S. and Saldanha-Da-Gama, F. (2010b) 'Single-assignment hub location problems with multiple capacity levels', *Transportation Research Part B: Methodological*, Vol. 44, No. 8, pp.1047–1066.

Cruijssen, F., Borm, P., Dullaert, W. and Hamers, H. (2010) 'A versatile framework for cooperative hub network development', *European Journal of Industrial Engineering*, Vol. 4, No. 2, pp.210–227.

de Camargo, R.S., de Miranda Jr., G. and Ferreira, R.P. (2011) 'A hybrid outer-approximation/ benders decomposition algorithm for the single allocation hub location problem under congestion', *Operations Research Letters*, Vol. 39, No. 5, pp.329–337.

de Sá, E.M.D., Morabito, R. and de Camargo, R.S. (2018) 'Benders decomposition applied to a robust multiple allocation incomplete hub location problem', *Computers & Operations Research*, Vol. 89, pp.31–50.

Ebery, J., Krishnamoorthy, M., Ernst, A. and Boland, N. (2000) 'The capacitated multiple allocation hub location problem: formulations and algorithms', *European Journal of Operational Research*, Vol. 120, No. 3, pp.614–631.

Elhedhli, S. and Wu, H. (2010) 'A Lagrangean heuristic for hub-and-spoke system design with capacity selection and congestion', *INFORMS Journal on Computing*, Vol. 22, No. 2, pp.282–296.

Ernst, A.T. and Krishnamoorthy, M. (1996) 'Efficient algorithms for the uncapacitated single allocation p-hub median problem', *Location Science*, Vol. 4, No. 3, pp.139–154.

Ernst, A.T. and Krishnamoorthy, M. (1998) 'Exact and heuristic algorithms for the uncapacitated multiple allocation p-hub median problem', *European Journal of Operational Research*, Vol. 104, No. 1, pp.100–112.

Ernst, A.T. and Krishnamoorthy, M. (1999) 'Solution algorithms for the capacitated single allocation hub location problem', *Annals of Operations Research*, Vol. 86, pp.141–159.

Farahani, R.Z., Hekmatfar, M., Arabani, A.B. and Nikbakhsh, E. (2013) 'Hub location problems: a review of models, classification, solution techniques, and applications', *Computers & Industrial Engineering*, Vol. 64, No. 4, pp.1096–1109.

Gelareh, S., Monemi, R.N. and Nickel, S. (2015) 'Multi-period hub location problems in transportation', *Transportation Research Part E: Logistics and Transportation Review*, Vol. 75, pp.67–94.

Ghaffarinasab, N., Motallebzadeh, A., Jabarzadeh, Y. and Kara, B.Y. (2018) 'Efficient simulated annealing based solution approaches to the competitive single and multiple allocation hub location problems', *Computers & Operations Research*, Vol. 90, pp.173–192.

Hernandez, S., Unnikrishnan, A. and Awale, S. (2011) 'A less-than-truckload centralized carrier collaboration multi-hub location problem: a hybrid hub-and-spoke network', *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2263, pp.20–28.

Hu, Q-M. (2017) 'A reconfiguration optimisation model for hub-and-spoke network mergers', *European Journal of Industrial Engineering*, Vol. 11, No. 1, pp.101–132.

Ishfaq, R. and Sox, C.R. (2012) 'Design of intermodal logistics networks with hub delays', *European Journal of Operational Research*, Vol. 220, No. 3, pp.629–641.

Kewcharoenwong, P. and Üster, H. (2017) 'Relay network design with capacity and link-imbalance considerations: a Lagrangean decomposition algorithm and analysis', *Transportation Science*, Vol. 51, No. 4, pp.1177–1195.

Kian, R. and Kargar, K. (2016) 'Comparison of the formulations for a hub-and-spoke network design problem under congestion', *Computers & Industrial Engineering*, Vol. 101, pp.504–512.

Kim, H. and O'Kelly, M.E. (2009) 'Reliable p-hub location problems in telecommunication networks', *Geographical Analysis*, Vol. 41, No. 3, pp.283–306.

Lin, C-C. and Lee, S-C. (2010) 'The competition game on hub network design', *Transportation Research Part B: Methodological*, Vol. 44, No. 4, pp.618–629.

Lin, C-C. and Lee, S-C. (2015) 'Zone pricing for time-definite LTL freight transportation with elastic demand', *Computers & Operations Research*, Vol. 62, pp.51–60.

Lüer-Villagra, A. and Marianov, V. (2013) 'A competitive hub location and pricing problem', *European Journal of Operational Research*, Vol. 231, No. 3, pp.734–744.

Mahmutogullari, A.I. and Kara, B.Y. (2016) 'Hub location under competition', *European Journal of Operational Research*, Vol. 250, No. 1, pp.214–225.

Mohammadi, M., Torabi, S.A. and Tavakkoli-Moghaddam, R. (2014) 'Sustainable hub location under mixed uncertainty', *Transportation Research Part E: Logistics and Transportation Review*, Vol. 62, pp.89–115.

Monemi, R.N., Gelareh, S., Nagih, A. and Jones, D. (2020) 'Bi-objective load balancing multiple allocation hub location: a compromise programming approach', *Annals of Operations Research* [online] https://xs.scihub.ltd/https://doi.org/10.1007/s10479-019-03421-w.

O'Kelly, M.E. (1987) 'A quadratic integer program for the location of interacting hub facilities', *European Journal of Operational Research*, Vol. 32, No. 3, pp.393–404.

O'Kelly, M.E. and Bryan, D.L. (1998) 'Hub location with flow economies of scale', *Transportation Research Part B: Methodological*, Vol. 32, No. 8, pp.605–616.

OR-Library (2015) *Hub Location* [online] http://people.brunel.ac.uk/~mastjjb/jeb/orlib/phubinfo.html (accessed 10 February 2020).

Sasaki, M. and Fukushima, M. (2003) 'On the hub-and-spoke model with arc capacity constraints', *Journal of the Operations Research Society of Japan-Keiei Kagaku*, Vol. 46, No. 4, pp.409–428.

Sasaki, M., Campbell, J.F., Krishnamoorthy, M. and Ernst, A.T. (2014) 'A Stackelberg hub arc location model for a competitive environment', *Computers & Operations Research*, Vol. 47, pp.27–41.

Yoon, M-G. and Current, J. (2008) 'The hub location and network design problem with fixed and variable arc costs: formulation and dual-based solution heuristic', *Journal of the Operational Research Society*, Vol. 59, No. 1, pp.80–89.