Profit Maximizing Hub Location Problems

Gita Taherkhani, Sibel A. Alumur

gita.taherkhani@uwaterloo.ca, sibel.alumur@uwaterloo.ca

Department of Management Sciences, University of Waterloo, Waterloo, Ontario, Canada

Abstract

In this paper, we study profit maximizing hub location problems. We formulate mathematical models determining the location of hubs, designing the hub networks, and routing the demand in order to maximize profit. The profit is calculated by summing the total revenue minus total cost. Total cost includes the total transportation cost, the installation cost of hubs, and the cost of operating hub links. We consider all possible allocation strategies: multiple allocation, single allocation, and r-allocation. As an extension, for each allocation strategy, we also model the cases in which direct connections between non-hub nodes are allowed. To test and evaluate the performances of the proposed models, we use two well-known data sets from the literature. We analyze the resulting hub networks under various different parameter settings.

Keywords: Hub location, hub network design, profit maximization.

1 Introduction

Hubs serve as switching, sorting, connecting, and consolidation points in many-to-many distribution networks. In hub networks, demand between origin-destination pairs are routed through hubs instead of using direct connections. The aim is to reduce the costs of establishing a network connecting many origins and destinations, and also to consolidate flows at hubs to exploit economies of scale. Hub networks have widespread applications in air passenger and freight transportation, express shipment, postal delivery, public transportation networks, and telecommunications (Campbell et al. [16] and Contreras [21]).

Hub location problems generally consist of two main decisions: the location of hubs and the allocation of demand nodes to these hubs. The optimal routes of flow on the hub network to satisfy the demand between origin-destination pairs is also to be determined. Several hub location problems are proposed in the literature. The reader may refer to reviews on this area by Campbell et al. [16], Alumur and Kara [3], Campbell and O'Kelly [19], and Contreras [21].

Hub location problems can be classified based on the design of the access network which represents the allocation of demand nodes to hubs. There are two basic allocation strategies in the literature: single allocation and multiple allocation. In single allocation, each node must be allocated to exactly one hub. In multiple allocation, on the other hand, there is no limit on the maximum number of hubs that a node can be allocated to. Additionally, Yaman [61] introduced the r-allocation strategy, where each node can be allocated to at most r hubs.

Different applications call for different design characteristics for access networks. For instance, in express delivery networks, non-hub nodes are customarily allocated to a single hub due to ease of management. However, single allocation may not be reasonable for other application areas, especially if multiple allocation strategy is economically justifiable. On the other hand, in some applications such as air passenger or public transportation, both single and multiple allocation strategies are two extremes. In such cases, r-allocation strategy can be considered where the number of connections between a non-hub node and a hub node is restricted by a fixed number. Additionally, allowing for the possibility of direct connections between non-hub nodes is also of interest, particularly in air passenger transportation where most of the clients prefer non-stop flights to arrive at their destinations or in freight transportation where there is enough demand to send direct shipments between non-hub nodes. Note that, in spite of the similarities between these allocation strategies, there are some fundamental differences which necessitate them to be modeled separately.

Classical hub location problems are modeled mainly from a cost point of view. Hence, generic models aim to minimize total network cost to satisfy all demand. However, from a profit point of view, it may be more advantageous not to serve all demand. As such, some origin-destination pairs may remain unserved when it is not profitable to be served. The decision on how much demand to serve should be made by considering the trade-off between revenue and cost.

In this paper, we focus on profit maximizing hub location problems. Taking profit into consideration, we allow for the possibility of serving a subset of the demand. In this regard, we assume that a portion of the demand can be unserved when it is not profitable. This assumption actually provides a more realistic framework for designing hub networks.

The profit maximizing hub location problems introduced in this study aim to find how many hubs to locate and where, how to allocate demand nodes to these hubs, the optimal design of the hub network, and optimal routes of flow to satisfy the demand selected to be served between origin-destination pairs.

Our contribution in this paper is to introduce new problems and develop mathematical models determining the locations of hubs, designing the hub networks, and routing the demand in order to maximize profit. We introduce novel mathematical formulations determining the

2

set of demand nodes to be served to maximize profit in hub location problems. We model the multiple and single allocation versions as well as the r-allocation strategy. For each allocation strategy, we also model the versions in which direct connections between non-hub nodes are allowed. We present extensive computational analysis to evaluate the solution potential of the proposed models with emphasis on the effect of the choice of parameters.

Potential applications of profit maximizing hub location problems arise in the design of airline passenger and freight transportation networks, truckload and less-than-truckload transportation, and express shipment and postal delivery. Our models are applicable to the design of new hub networks as well as for improving the existing ones. The proposed models can be used as decision support tools for evaluating prospective locations of hubs and deciding on the optimal shipment strategies, while determining the set of origin-destination pairs to be served.

The remainder of the paper is organized as follows: In the next section, we present a survey of the relevant literature. In Section 3, we define the problem and introduce some notation to be used in the following sections. We formulate mixed-integer linear programming models for the profit maximizing hub location problems in Section 4. In Section 5, we test and evaluate our models on instances derived from the CAB and AP datasets. Finally, Section 6 provides some concluding remarks.

2   Literature Review

As mentioned in the introduction, hub location problems presented in the literature are commonly classified according to the selection of allocation decisions. The proposed mathematical models and algorithms are usually dependent on these access network design decisions. For some example studies, the reader may refer to Ernst and Krishnamoorthy [29], Labbé et al. [36], and Chen [20] for single allocation, and to Ernst and Krishnamoorthy [30], Boland et al. [13], and Contreras et al. [24] for multiple allocation models and algorithms. There are studies simultaneously addressing both allocation strategies as well such as Campbell [14], Skorin-Kapov et al. [56], and O'Kelly et al. [49].

Regarding the design of the network in-between hubs, earlier studies on hub location assumed the inter-hub network to be complete with a direct link between every hub pair (for example, the seminal papers O'Kelly [45] and O'Kelly [46], Campbell [14], Ernst and Krishnamoorthy [29], and Skorin-Kapov et al. [56]). For many studies, this is a direct consequence of having triangle inequality for transportation costs and not having any fixed costs for inter-hub links in the models. However, many real-life hub networks from various applications do not operate fully interconnected hub networks. Despite significant contributions on designing complete inter-hub

networks, relatively little effort has been devoted to the possibility of designing incomplete inter-hub networks. Examples of such studies are Nickel et al. [44], Campbell et al. [17], Alumur et al. [5], Contreras et al. [25], Alumur et al. [6], and Lüer-Villagra and Marianov [37].

Designing distinct topologies for inter-hub networks are also considered in the literature. Contreras et al. [25], and Martins de Sá et al. [41] studied the problem in which hubs are connected by means of a tree network. Potential applications arise when the set-up costs for hub links are high. Yaman [60] designed a hierarchical three-level hub network, where the top level consists of a complete network and the second and third levels are unions of star networks. Alumur et al. [4] studied the design of a hierarchical star-mesh-star hub network with multiple transportation modes. Yaman and Elloumi [62] modeled the design of two-level star networks taking service quality considerations into account. Contreras et al. [26] analyzed the case where the hubs are connected by means of a cycle, and Martins de Sá et al. [42] by means of a line.

In most of the hub location studies, direct connections between non-hub nodes is not possible. Some authors included this possibility in their models such as Aykin [9], Aykin [10], Nickel et al. [44], Sung and Jin [57], Wagner [59], and Mahmutoğulları and Kara [38].

Other recent developments in hub location include modeling uncertainty (Contreras et al. [23], Alumur et al. [7], and Meraklı and Yaman [43]), incorporating the time dimension (Contreras et al. [22], Alumur et al. [8], and Correia et al. [27]), taking multiple modes of transportation into account (Alumur et al. [6] and Serper and Alumur [55]), and modeling reliability of hubs (Kim and O'Kelly [34], Azizi et al. [11], and Tran et al. [58]).

Most of the classical hub location models assume that the transportation cost of flows via hub links is reduced by an exogenous discount factor independent of flow (usually referred as α). However, as noted by Campbell and O'Kelly [19], taking a flow-independent discount for transportation on hub links may undermine the basic definition for economies of scale. To deal with this issue, different approaches have been considered in the literature, including flow-dependent discounts for inter-hub links using a non-linear cost function (O'Kelly and Bryan [48] and Horner and O'Kelly [32]), considering economies of scale with flow thresholds (Podnar et al. [52]), hub arc location models (Campbell et al. [17] and Campbell et al. [18]), incomplete inter-hub network designs (Alumur et al. [5] and Contreras et al. [25]), and models with fixed and variable cost components (Kimms [35], O'Kelly et al. [50], and Serper and Alumur [55]).

In this paper, we follow the classical assumption, where the economies of scale is exploited only between hubs and calculated by using a constant discount factor. Due to the complexity of our problems, as noted in O'Kelly and Bryan [48], it would be impractical to incorporate a flow-dependent economies of scale function directly into our models, which would lead to a non-linear cost function. On the other hand, as noted in Campbell [15], allowing incomplete hub networks

that do not require full connectivity of all hub nodes provides a better match between optimal flows and transportation cost discounts. This is exactly the case in our problem. Moreover, we analyze the resulting flows on network connections as well as the impact of flow-dependent economies of scale on transportation cost through our computational analysis to justify the use of a constant economies of scale discount factor.

There are not many studies in the literature focusing on maximization objectives within hub location problems. Perhaps, one of the early works is Campbell [14] introducing the hub maximal covering location models. Given a number of hub facilities to locate, the problem aims to maximize the demand covered. Campbell [14] defined different notions for coverage. Hwang and Lee [33] studied uncapacitated single allocation p-hub maximal covering problem and proposed a heuristic algorithm for the problem. More recently, Peker and Kara [51] extended the definition of coverage and introduced the notion of partial coverage that changes with distance. They developed mixed-integer programming formulations with partial coverage for single and multiple allocation versions.

Similar to hub maximal covering problems, we do not force all demand nodes to be served in this study. However, we do not have a given budget for locating hubs and we maximize profit instead of the covered demand. Moreover, we also consider total transportation cost which is not taken into account in covering type hub location studies.

There are studies determining the locations of hubs within a competitive environment. In these studies, instead of having a single firm, there are a number of firms competing to serve the demand. Hence, competitors' decisions affect the profit of a firm. Different objective functions are considered in competitive hub location problems, such as maximizing the demand captured, and maximizing total revenue or profit. Examples of studies considering a competitive environment include Marianov et al. [39], Eiselt and Marianov [28], Gelareh et al. [31], and Lüer-Villagra and Marianov [37]. There are also some studies considering a game theoretic framework in addressing competitive hub location problems such as Sasaki and Fukushima [54] and Sasaki et al. [53].

Even though objective functions of some competitive hub location studies also aim at maximizing profit, we do not consider a competitive environment in this study. In our study, there is only a single firm that wants to design its hub network in the most profitable way. Moreover, we do not force all demand to be served.

Alibeyg et al. [1] introduced a hub network design problem with profits. The problem aims to determine the locations of hubs, decide which edges to activate, select pairs of nodes and a set of commodities to be served, and make routing decisions with the objective of maximizing total profits. They considered a multiple allocation setting and assumed that each origin and

destination path consists of at least one and at most two hub nodes. The demand between two pairs of nodes is served through at most three edges. They modeled different variations of the problem and use CAB dataset to test the performance of their models using CPLEX. In a subsequent study, Alibeyg et al. [2] proposed an exact algorithm for the profit-oriented models introduced by Alibeyg et al. [1]. They embedded Lagrangean relaxation within a branch-and-bound algorithm. They also used reduction tests and partial enumeration to reduce the problem size as well as the computational effort.

Perhaps the most related study to ours is Alibeyg et al. [1]. Similar to Alibeyg et al. [1], we study hub network design problems with profits and introduce profit maximizing objective functions. In this paper, we provide a new modeling framework for the problem. Unlike Alibeyg et al. [1], we do not assume that the demand between pair of nodes can be served through at most two hubs or three edges. In our problem, the demand of an origin-destination pair can be shipped through any number of hubs and network connections as necessary. We include additional design variables to account for the design of the inter-hub network. Incomplete hub networks are employed in many real-life applications including but not limited to freight transportation, rapid transit systems, express shipment and postal delivery networks. Therefore, relaxing a commonly used assumption on the structure of the origin-destination paths, which limits the number of hub nodes to two, leads to a more realistic model where non-trivial routing decisions are explicitly taken into account. In our computational analysis, we present a direct comparison of our results with one of the models presented in Alibeyg et al. [1]. Furthermore, in this paper, we study multiple, single, and r-allocation versions of the problem and also allow for the possibility of direct connections between non-hub nodes.

3   Problem Definition and Notation

The hub location problems that we study in this paper determine the location of hubs and design the hub network in order to maximize profit. The location decision focuses on the selection of a set of nodes to establish hubs, whereas the network design decisions deal with the design of the links to connect nodes of the network. The optimal routes of flows through the network are also to be determined.

There is a given set of nodes as well as the demand that can be served between them. Demand is defined between pairs of nodes. The problem is to determine which set of origin-destination (O-D) pairs to serve and how to serve them to maximize profit. Taking profit into consideration, the problem measures the trade-off between the revenue and cost in determining which O-D pairs to serve. In this regard, it is assumed that some O-D pairs can remain unserved

6

if it is not profitable to serve them.

We assume that hubs have enough capacity to handle all flow. In this case, the demand of an O-D pair can be satisfied completely, if it is profitable and partial satisfaction of the demand would be unreasonable. The amount of demand handled at hubs may in turn effect the capacities of the hubs to be established. However, we do not consider the capacities of the hubs to be established in our models. The capacity of a hub is dependent on various factors including decisions related to frequency and scheduling. For example, the capacity of an airport hub is dependent on the maximum hourly or daily numbers of aircraft movements. Similarly, the capacity of a transfer hub in an express shipment delivery network depends on the amount of parcels that can be sorted, for example, in an hour. As we do not model the time dimension in this study, it is difficult to assess and account for the capacity of a hub.

A set of hub nodes are to be located to serve the demand. All nodes can be potential hub facilities. A set of hub links to operate between hub nodes is also to be determined. We do not impose a fully interconnected hub network. Demand of an O-D pair can be shipped on any number of network connections as necessary. Moreover, no particular network structure (such as a star or a tree network) is assumed for the inter-hub network.

For the design of the access network, we consider all possible allocation strategies introduced in the literature; i.e., single allocation, multiple allocation, and r-allocation. Additionally, we consider the cases where the demand of an O-D pair can be shipped directly without using any intermediary hub facilities. Considering all the possibilities for shipments, the problem is to design the hub network in the most profitable way.

The objective of our problem is the maximization of total profit. Total profit is calculated by subtracting total cost from the total revenue. Revenue is obtained from satisfying the demand of each O-D pair. We assume that price is exogenously determined by the market and that the revenue is independent from the route of flow of the demand.

Total cost includes the variable transportation cost, the fixed installation cost of hubs, and the fixed cost of activating links. The transportation cost on the hub network between each O-D pair is calculated by the cost of transportation from origin-to-hub (collection), between hubs (transfer), and from hub-to-destination (distribution). In case of direct connections, there is no collection, transfer, and distribution legs. In this case, transportation cost is calculated by the cost of shipping all demand between an O-D pair using the direct connection in-between. Unit transportation costs do not need to satisfy triangle inequality. In addition to transportation cost, there is a fixed cost involved for operating an inter-hub link as well as a direct link between non-hub nodes. No operational cost is considered for the allocation connections.

There is economies of scale between hubs, for example, due to bulk transportation or frequent

service. It is assumed that economies of scale is reflected on the transportation cost by a constant discount factor. We provide extensive analysis with varying values of the economies of scale factor.

We introduce the following notation for the parameters required for modeling the problem:

N	Set of nodes.
w_{ij}	Amount of demand to be shipped from node $i \in N$ to node $j \in N$.
	($O_i = \sum\limits_{j \in N} w_{ij}$ denotes the flow originated at node $i \in N$.)
r_{ij}	Revenue from satisfying a unit demand from node $i \in N$ to node $j \in N$.
c_{ij}	Unit cost of transportation from node $i \in N$ to node $j \in N$.
f_k	Fixed installation cost of a hub at node $k \in N$.
g_{kl}	Fixed cost of activating a hub link from hub $k \in N$ to hub $l \in N$.
q_{ij}	Fixed cost of activating a direct link from node $i \in N$ to node $j \in N$.
α	Transportation cost discount factor ($0 \le \alpha < 1$).

We model single, multiple, and r-allocation versions of the profit maximizing hub location problem. For each allocation rule, we also model the cases in which direct connections between non-hub nodes are allowed. In the following section, we introduce the mathematical formulations for different versions of the problem.

4 Mathematical Formulations

In this section, we present mixed-integer linear programming (MILP) models for different versions of the profit maximizing hub location problem. In the next section, we introduce our MILP model for the multiple allocation problem. In Section 4.2, we present MILP models for the single allocation and the r-allocation versions of the problem. In Section 4.3, we introduce models allowing for direct connections with all allocation rules.

4.1 Multiple allocation model

We now introduce a mathematical model for the multiple allocation version of the problem. In this problem, a non-hub node can be allocated to as many hubs as it is profitable. So, there is no limit on the number of hubs that a non-hub node can be allocated to. The decision variables that need to be defined for the multiple allocation model are as follows:

$$h_k = \begin{cases} 1, & \text{if a hub is located at node } k \in N, \\ 0, & \text{otherwise.} \end{cases}$$

$$y_{ijkl} = \begin{cases} 1, & \text{if the demand between node } i \in N \text{ and } j \in N \text{ is satisfied through a path} \\ & \text{with the first hub } k \in N \text{ and the last hub } l \in N, \\ 0, & \text{otherwise.} \end{cases}$$

$$z_{kl} = \begin{cases} 1, & \text{if an inter-hub link is operating from hub } k \in N \text{ to hub } l \in N, \\ 0, & \text{otherwise.} \end{cases}$$

f_{ikl} = Amount of demand originated at node $i \in N$ and routed on the inter-hub link from hub $k \in N$ to hub $l \in N$.

The *profit maximizing multiple allocation hub location problem* is modeled as:

$$\text{Max} \quad \sum_{i \in N}\sum_{j \in N}\sum_{k \in N}\sum_{l \in N} r_{ij}w_{ij}y_{ijkl} - \Big[\sum_{i \in N}\sum_{j \in N}\sum_{k \in N}\sum_{l \in N}(c_{ik}+c_{lj})w_{ij}y_{ijkl} + \sum_{i \in N}\sum_{k \in N}\sum_{l \in N}\alpha c_{kl}f_{ikl}$$

$$+ \sum_{k \in N} f_k h_k + \sum_{k \in N}\sum_{l \in N} g_{kl}z_{kl} \Big] \tag{1}$$

$$\text{s.t.} \quad \sum_{k \in N}\sum_{l \in N} y_{ijkl} \le 1 \qquad\qquad\qquad\qquad i,j \in N \tag{2}$$

$$\sum_{l \in N} y_{ijkl} + \sum_{l \in N: l \neq k} y_{ijlk} \le h_k \qquad\qquad i,j,k \in N \tag{3}$$

$$\sum_{l \in N, l \neq k} f_{ikl} - \sum_{l \in N, l \neq k} f_{ilk} = \sum_{j \in N}\sum_{l \in N} w_{ij}y_{ijkl} - \sum_{j \in N}\sum_{l \in N} w_{ij}y_{ijlk} \qquad i,k \in N \tag{4}$$

$$f_{ikl} \le O_i z_{kl} \qquad\qquad\qquad\qquad\qquad i,k,l \in N, k \neq l \tag{5}$$

$$z_{kl} \le h_k \qquad\qquad\qquad\qquad\qquad\qquad k,l \in N, k \neq l \tag{6}$$

$$z_{kl} \le h_l \qquad\qquad\qquad\qquad\qquad\qquad k,l \in N, k \neq l \tag{7}$$

$$f_{ikl} \ge 0 \qquad\qquad\qquad\qquad\qquad\qquad i,k,l \in N \tag{8}$$

$$h_k \in \{0,1\} \qquad\qquad\qquad\qquad\qquad k \in N \tag{9}$$

$$y_{ijkl} \in \{0,1\} \qquad\qquad\qquad\qquad\qquad i,j,k,l \in N \tag{10}$$

$$z_{kl} \in \{0,1\} \qquad\qquad\qquad\qquad\qquad k,l \in N, k \neq l \tag{11}$$

The objective function value (1) represents net profit. Total cost is subtracted from the total revenue to calculate the net profit. The first term of the objective function calculates the revenue obtained from satisfying the demand. The terms in parenthesis represent the transportation cost, the installation cost of hubs, and the cost of operating hub links, respectively. While calculating

the total transportation cost, the hub-to-hub transportation is discounted by the economies of scale discount factor, α.

Constraints (2) state that there must be a unique path recognized by the first and the last hubs, if the demand between a given pair of nodes is to be satisfied. Constraints (3) ensure that the demand between origin and destination nodes can be satisfied only through located hubs. Constraints (4) are the flow balance equations. Constraints (5) enforce that the flow is routed only on the operated hub links. Constraints (6) and (7) indicate that an inter-hub link can only be operated if both of the end nodes of that link are hubs. Constraints (8)-(11) represent the non-negative and binary variables.

In an uncapacitated environment, there always exists an optimal solution of the problem where the demand of each O-D pair is either satisfied fully or not satisfied at all. Therefore, we defined y_{ijkl} variables as binary. However, integrality property holds for these variables. Hence, even though if we let $y_{ijkl} \geq 0$ for all $i, j, k, l \in N$, there exists an optimal solution of this model where $y_{ijkl} \in \{0, 1\} \, \forall \, i, j, k, l \in N$. In our computational experiments, we solved the model with y_{ijkl} variables being binary as we obtained better solution times compared with real variables.

As mentioned in the problem definition, we assume that hubs have enough capacity to handle all flow. If this is not the case, then one may add the following sets of constraints into the model:

$$\sum_{i \in N} \sum_{l \in N} f_{ikl} \leq \Gamma_k h_k \qquad k \in N \tag{12}$$

$$y_{ijkl} \geq 0 \qquad\qquad i, j, k, l \in N \tag{13}$$

where Γ_k is defined as the available capacity of a hub located at node $k \in N$. In the presence of capacity constraints (12), y_{ijkl} variables should be defined as continuous routing variables denoting the fraction of the demand from node $i \in N$ to $j \in N$ that is satisfied through a path with the first hub $k \in N$ and the last hub $l \in N$.

4.2 Single allocation and r-allocation models

In this section, we first introduce the mathematical model for the single allocation profit maximizing hub location problem. In the single allocation problem, each non-hub node can be allocated to at most one hub node. To model this problem, in addition to the decision variables introduced in the previous section, we define an additional binary variable as follows:

$$x_{ik} = \begin{cases} 1, & \text{if demand node } i \in N \text{ is allocated to hub node } k \in N, \\ 0, & \text{otherwise.} \end{cases}$$

($x_{kk} = 1$ indicates that a hub is located at node $k \in N$.)

The rest of the variables are the same as introduced in the multiple allocation model. The *profit maximizing single allocation hub location problem* is modeled as:

$$\text{Max} \quad \sum_{i \in N} \sum_{j \in N} \sum_{k \in N} \sum_{l \in N} r_{ij} w_{ij} y_{ijkl} - \Big[\sum_{i \in N} \sum_{j \in N} \sum_{k \in N} \sum_{l \in N} (c_{ik} + c_{lj}) w_{ij} y_{ijkl} + \sum_{i \in N} \sum_{k \in N} \sum_{l \in N} \alpha c_{kl} f_{ikl}$$

$$+ \sum_{k \in N} f_k x_{kk} + \sum_{k \in N} \sum_{l \in N} g_{kl} z_{kl} \Big] \tag{14}$$

s.t. $(2), (4), (5), (8), (10), (11)$

$$\sum_{k \in N} x_{ik} \leq 1 \qquad\qquad\qquad i \in N \tag{15}$$

$$x_{ik} \leq x_{kk} \qquad\qquad\qquad i, k \in N \tag{16}$$

$$y_{ijkl} \leq x_{ik} \qquad\qquad\qquad i, j, k, l \in N \tag{17}$$

$$y_{ijkl} \leq x_{jl} \qquad\qquad\qquad i, j, k, l \in N \tag{18}$$

$$z_{kl} \leq x_{kk} \qquad\qquad\qquad k, l \in N, k \neq l \tag{19}$$

$$z_{kl} \leq x_{ll} \qquad\qquad\qquad k, l \in N, k \neq l \tag{20}$$

$$x_{ik} \in \{0, 1\} \qquad\qquad\qquad i, k \in N \tag{21}$$

The objective function value (14) accounts for the net profit obtained from summing the total revenue minus total cost as in the multiple allocation version. Constraints (15) ensure that every demand node can be allocated to at most one hub node. To guarantee that the demand nodes can be allocated to only installed hubs, constraints (16) are included in the model. Constraints (17) and (18) link path variables with allocation decisions. Constraints (19) and (20) ensure that an inter-hub link is operated only in-between hubs. Constraints (21) are the domain constraints. The rest of the constraints are the same as introduced in the multiple allocation model.

As mentioned in Marin et al. [40], constraints (17) and (18) can be strengthened by replacing them with the following sets of constraints:

$$\sum_{l \in N} y_{ijkl} \leq x_{ik} \qquad i, j, k \in N \tag{22}$$

$$\sum_{k \in N} y_{ijkl} \leq x_{jl} \qquad i, j, l \in N \tag{23}$$

Even though using the above strengthened constraints will result in a better LP relaxation bound, in our preliminary experiments, we observed that the solution times of the model using a commercial solver gets worse. Thus, we employed constraints (17) and (18) in our computational experiments.

Yaman [61] introduced the r-allocation version of a hub location problem where each node can be allocated to at most r hubs. This version is actually a generalization of the single and multiple allocation versions. When $r = 1$, the problem reduces to single allocation, whereas, when $r = |N|$, it reduces to multiple allocation. In order to model the r-allocation version of the profit maximizing hub location problem, we introduce the following set of constraints to replace constraints (15) in the single allocation model:

$$\sum_{k \in N} x_{ik} \leq r \qquad i \in N \tag{24}$$

The MILP formulation of the *profit maximizing r-allocation hub location problem* consists of the objective function (14) and constraints (2), (4), (5), (8), (10), (11), (16)-(21), and (24).

4.3   Models allowing for direct connections

In the previous models, we assumed that direct transportation between demand nodes (without using any hubs) is not allowed. In this section, we model the variation in which direct connections between non-hub nodes are allowed. To allow for direct services, we need to define an additional binary variable s_{ij} as follows:

$$s_{ij} = \begin{cases} 1, & \text{if there is a direct connection from non-hub node } i \in N \text{ to non-hub node } j \in N, \\ 0, & \text{otherwise.} \end{cases}$$

We first introduce the *profit maximizing multiple allocation hub location problem with direct connections*:

$$
\begin{aligned}
\text{Max} \quad & \sum_{i \in N} \sum_{j \in N} \sum_{k \in N} \sum_{l \in N} r_{ij} w_{ij} y_{ijkl} + \sum_{i \in N} \sum_{j \in N} r_{ij} w_{ij} s_{ij} - \Bigg[\sum_{i \in N} \sum_{j \in N} \sum_{k \in N} \sum_{l \in N} (c_{ik} + c_{jl}) w_{ij} y_{ijkl} \\
& + \sum_{i \in N} \sum_{k \in N} \sum_{l \in N} \alpha c_{kl} f_{ikl} + \sum_{i \in N} \sum_{j \in N} w_{ij} c_{ij} s_{ij} + \sum_{k \in N} f_k h_k + \sum_{k \in N} \sum_{l \in N} g_{kl} z_{kl} \\
& + \sum_{i \in N} \sum_{j \in N} q_{ij} s_{ij} \Bigg]
\end{aligned}
\tag{25}
$$

$$
\begin{aligned}
\text{s.t.} \quad & (3) - (11) \\
& s_{ij} + h_j \leq 1 & i, j \in N & \tag{26} \\
& s_{ij} + h_i \leq 1 & i, j \in N & \tag{27} \\
& s_{ij} + \sum_{k \in N} \sum_{l \in N} y_{ijkl} \leq 1 & i, j \in N & \tag{28} \\
& s_{ij} \in \{0, 1\} & i, j \in N & \tag{29}
\end{aligned}
$$

12

The objective function (25) sums the revenues obtained from satisfying the demand through the hub network and also through direct connections. In addition to the previously defined costs, the total cost includes the cost of operating direct links between non-hub nodes as well. To guarantee that the direct connections are operated only between non-hub nodes, constraints (26) and (27) are included in the model. Constraints (28) result from modifying constraints (2) to ensure that there is either a unique path through the hub network or a direct link to satisfy the demand between a given pair of nodes. Finally, constraints (29) provide the domain for the binary direct-connection variables.

We also model the *profit maximizing single allocation hub location problem with direct connections*:

$$
\begin{aligned}
\text{Max} \quad & \sum_{i \in N} \sum_{j \in N} \sum_{k \in N} \sum_{l \in N} r_{ij} w_{ij} y_{ijkl} + \sum_{i \in N} \sum_{j \in N} r_{ij} w_{ij} s_{ij} - \Big[\sum_{i \in N} \sum_{j \in N} \sum_{k \in N} \sum_{l \in N} (c_{ik} + c_{jl}) w_{ij} y_{ijkl} \\
& + \sum_{i \in N} \sum_{k \in N} \sum_{l \in N} \alpha c_{kl} f_{ikl} + \sum_{i \in N} \sum_{j \in N} w_{ij} c_{ij} s_{ij} + \sum_{k \in N} f_k x_{kk} + \sum_{k \in N} \sum_{l \in N} g_{kl} z_{kl} \\
& + \sum_{i \in N} \sum_{j \in N} q_{ij} s_{ij} \Big]
\end{aligned}
\tag{30}
$$

$$
\text{s.t.} \quad (4), (5), (8), (10), (11), (15) - (21), (28), (29)
$$

$$
s_{ij} + x_{jj} \leq 1 \qquad\qquad\qquad\qquad\qquad\qquad i, j \in N \tag{31}
$$

$$
s_{ij} + x_{ii} \leq 1 \qquad\qquad\qquad\qquad\qquad\qquad i, j \in N \tag{32}
$$

The objective function (30) calculates net profit considering the revenue obtained through direct connections as well as the cost of operating those links. Constraints (31) and (32) are included in the model to assure that direct connections are allowed only between non-hub nodes. The rest of the constraints are the same as introduced in the previous models.

For the r-allocation model with direct connections, we simply need to replace constraint (15) with constraint (24) in the single allocation version of the formulation.

4.4 Variable Fixing

To decrease the computational burden with all the proposed models, the values of some decision variables can be fixed with preprocessing. To this end, we exploit the following properties of our mathematical models:

$$
y_{ijkl} = 0 \qquad i, j, k, l \in N : r_{ij} < (c_{ik} + c_{lj}) \tag{33}
$$

13

As noted in the problem definition, only profitable demand will be served. Hence, when the revenue from satisfying a unit demand between O-D pair $i, j \in N$ is strictly smaller than the sum of unit transportation costs from origin i to hub k (collection) and from hub l to destination j (distribution), no profit can be obtained from satisfying the demand of the O-D pair $i, j \in N$ through hubs $k, l \in N$. Accordingly, the optimal value for the variable y_{ijkl} can be set to zero when the above mentioned condition is met.

In the next section, we present computational analysis with all of the introduced mathematical formulations.

5 Computational Results

We performed extensive computational experiments to analyze the resulting hub networks and the performance of the proposed mathematical models. We used two well-known benchmark data sets from the literature for hub location: the U.S. Civil Aeronautics Board (CAB) and the Australia Post (AP) datasets. Both of these datasets are readily available in OR Library [12].

Computational experiments were carried out on a workstation that contains: Intel Core i7-3930K 2.61GHz CPU, and 39 GB of RAM. The mathematical models were solved using IBM ILOG CPLEX 12.7. Variable fixing was employed as detailed in Section 4.4. All the instances were solved to optimality (10^{-5} gap) using the default settings.

The organization of this section is as follows: In the next section, we present computational results with the CAB dataset. In this section, we also provide some managerial insights obtained from solving the models on this data set. Section 5.2, presents analysis on the economies of scale factor. In Section 5.3, we compare our results with results from the relevant literature. To better understand the performance of the mathematical models from a computational point of view, we present numerical results with larger-sized instances derived from the AP dataset in Section 5.4.

5.1 Results with the CAB dataset

CAB dataset is based on airline passenger interactions between 25 cities in United States in 1970 (O'Kelly [46]). OR Library provides the transportation costs (c_{ij}) and the demand between each pair of cities (w_{ij}) for this dataset (Beasley [12]). As customarily done in the literature, we scaled the demand values so that the total demand adds up to one. Since CAB dataset does not provide any information regarding the revenues, we used average transportation costs to estimate values for the revenues. We test three different values such that $r_{ij} \in \{1000, 1500, 2000\}$ for each O-D pair $i, j \in N$, referred to as low, medium, and high revenue levels, respectively. Motivated from

the postal delivery applications, where the price of sending a letter is the same between any two locations, the revenue per unit demand is taken to be independent from the O-D pair. We also test three different values for hub installation costs generated by O'Kelly [47] such that $f_k \in \{50, 100, 150\}$ for all $k \in N$. Similarly, we refer these values as low, medium, and high cost levels, respectively. Cost of operating an inter-hub link is set to be 10% of hub installation costs; i.e., $g_{kl} = 0.1f_k$ for all $k, l \in N$. The cost of operating a direct link, on the other hand, is set to be the 20% of the operational cost of an inter-hub link; i.e., $q_{ij} = 0.2g_{ij}$ for all $i, j \in N$. These values can be interpreted by assuming that the frequency of service on the inter-hub links is five times more than that of direct connections. Operational costs are dependent on the level of hub installations costs (low, medium, or high) and taken to be the same for all potential links. The discount factor α is taken as 0.2, 0.4, 0.6, and 0.8 as customarily done in the literature. Table 1 summarizes the parameter settings with the CAB dataset.

Table 1: Parameter settings with the CAB dataset.

Description	Parameter	Value
Set of nodes	$|N|$	25
Demands	w_{ij}	OR Library [12]
Revenue per unit demand	r_{ij}	Low: 1000, Medium: 1500, High: 2000
Transportation cost per unit of flow	c_{ij}	OR Library [12]
Installation cost of a hub	f_k	Low: 50, Medium: 100, High: 150
Operational cost of an inter-hub link	g_{kl}	Low: 5, Medium: 10, High: 15
Operational cost of a direct link	q_{ij}	Low: 1, Medium: 2, High: 3
Discount factor for inter-hub connections	α	0.2, 0.4, 0.6, 0.8

We initially took runs with the multiple and single allocation models (without any direct connections) under different parameter values. The results obtained from solving the models are summarized in Table 2. The first two columns report the values of the cost parameters and the discount factor, respectively. For each allocation rule, the next three columns indicate the maximum net profit, total percentage of satisfied demand, and the locations of the hub nodes, respectively, in the optimal solutions of the corresponding instances. The "CPU time (s)" columns present the run time of instances (in seconds) obtained from solving the problems to optimality. To demonstrate the results with different revenue levels, Table 2 is split horizontally into three parts for high, medium, and low revenue levels.

Observe from Table 2 that when the costs increase, net profit values decrease along with the percentage of satisfied demand. Note that net profit is directly proportional to the percent of satisfied demand. With increased cost levels, the models tend to result in locating fewer hubs. As expected, when the number of opened hubs in the solutions increase, CPU time required for solving the instances to optimality also increase. Hence, low cost instances tend to require more CPU time compared with high cost instances.

Table 2: Multiple and single allocation solutions with the CAB dataset.

Costs	α	Multiple Allocation				Single Allocation			
		Net profit	Satisfied demand (%)	Hub locations	CPU time (s)	Net profit	Satisfied demand (%)	Hub locations	CPU time (s)
					High revenue				
Low	0.2	1190	100.00	4,7,12,14,17	4891	1181	99.67	4,7,12,14,17	12343
	0.4	1033	97.33	4,7,12,14,17	4130	1013	96.33	4,7,12,14,17	9384
	0.6	926	94.00	4,7,12,14,17	2428	874	83.50	4,12,17	8933
	0.8	850	86.83	1,4,12,17	754	774	70.67	4,12,18	9048
Medium	0.2	913	99.00	4,12,17,24	802	895	96.33	4,12,17	8280
	0.4	804	93.00	4,12,17	615	790	92.00	4,12,17	7271
	0.6	735	89.33	4,12,17	520	690	82.33	4,12,18	8138
	0.8	681	79.67	18,21	57	649	67.00	20	4102
High	0.2	740	97.00	4,12,17	550	735	96.33	4,12,17	8738
	0.4	634	93.00	4,12,17	362	620	92.00	4,12,17	6490
	0.6	599	67.00	20	90	599	67.00	20	1363
	0.8	599	67.00	20	28	599	67.00	20	1324
					Medium revenue				
Low	0.2	690	94.00	4,7,12,14,17	1756	685	93.33	4,7,12,14,17	9116
	0.4	550	89.00	4,7,12,14,17	1249	532	86.33	4,7,12,14,17	7924
	0.6	456	68.00	4,12,17	504	423	55.67	4,18	7180
	0.8	426	64.00	4,12,17	283	371	51.67	4,18	5597
Medium	0.2	429	83.17	4,12,17	549	418	81.67	4,12,17	4329
	0.4	349	62.67	4,17	142	331	58.67	4,17	2854
	0.6	328	61.00	4,17	33	310	52.67	20	1546
	0.8	325	61.00	4,18	23	310	52.67	20	1388
High	0.2	267	62.67	4,17	214	262	62.00	4,17	3414
	0.4	260	52.67	20	64	260	52.67	20	1100
	0.6	260	52.67	20	19	260	52.67	20	1073
	0.8	260	52.67	20	17	260	52.67	20	1060
					Low revenue				
Low	0.2	199	58.83	4,12,14,17	80	192	57.33	4,12,14,17	6057
	0.4	157	35.33	4,17	39	146	33.67	4,17	1152
	0.6	142	33.33	4,17	11	119	27.33	4,17	1031
	0.8	132	31.33	4,17	9	115	16.00	17	1121
Medium	0.2	69	36.00	4,17	21	67	36.00	4,17	939
	0.4	65	16.00	17	13	65	16.00	17	802
	0.6	65	16.00	17	10	65	16.00	17	789
	0.8	65	16.00	17	7	65	16.00	17	895
High	0.2	15	16.00	17	7	15	16.00	17	622
	0.4	15	16.00	17	6	15	16.00	17	584
	0.6	15	16.00	17	6	15	16.00	17	591
	0.8	15	16.00	17	5	15	16.00	17	421

We next observe the effect of the economies of scale factor on the solutions presented in Table 2. When the α value increases; that is, when the effect of economies of scale on transportation costs is lower, the net profit values and the percentage of satisfied demand decrease. The number of located hubs, on the other hand, either decreases or remains the same. Note that, in Table 2, there are some instances with high cost levels resulted in locating only one hub. Economies of scale factor has no effect in such instances. For a given revenue and cost level, the instances resulted in opening one hub yield the same net profit value and the same percentage of satisfied

demand independent from the value of α.

We tested all instances with three different revenue levels. When the revenue level decreases from high to low, the net profit values as well as the percentages of satisfied demand decrease. For a given cost level and α value, an increase in the revenue level results in satisfying more demand and may require opening more hubs.

Multiple allocation problem provides an upper bound for the single allocation problem in terms of net profit and percentage of satisfied demand. From the computational point of view, the single allocation model is more time consuming than the multiple allocation counterpart. In addition to the variables introduced for the multiple allocation model, recall that an additional binary variable was defined to model the single allocation problem. Hence, single allocation model is expected to be more challenging than the multiple allocation version.

The most time-consuming instance with the single allocation problem lasted around 3.43 hours, whereas the longest instance took 1.36 hours with the multiple allocation model. High revenue instances required more CPU time than medium and low revenue ones. This is because more hubs were opened with higher revenue levels to satisfy more demand. In general, the average CPU time requirements of the instances listed in Table 2 were 9.23 minutes and 1.13 hours for multiple and single allocation problems, respectively.

We also calculated the linear programming (LP) relaxation gaps of the multiple and single allocation models with all the instances listed in Table 2. Most of the gaps are strictly positive where the average LP gaps for the multiple and single allocation models are 4.46% and 15.87%, respectively. The LP gaps vary a lot with different parameter values. The maximum LP gap that we observed among the instances in Table 2 was 16.45% with the multiple allocation and 28.02% with the single allocation models, whereas the minimum was 0.02% with the multiple allocation and 7.24% with the single allocation model.

With both of the models, the locations of the hubs and also the corresponding network designs differ a lot depending on the revenue and cost levels, and the α value. However, there are some common nodes that are selected as hubs in most of the instances, such as Chicago (4), Los Angeles (12), and New York (17). This is because these nodes generate higher amount of demand compared with other nodes in the CAB dataset.

We depict the optimal networks of a few instances from the multiple and single allocation solutions in Figures 1 and 2. In both of these figures, squares represent the established hubs, bold lines the inter-hub links, and the thin lines the allocation connections. Since CAB dataset is symmetric, hub arcs are activated for both directions. Hence, the depicted links are undirected. There are some demand nodes in these figures without any connections at all (see for example Figures 1d and 2d). In such cases, no demand generated from those nodes is served in the
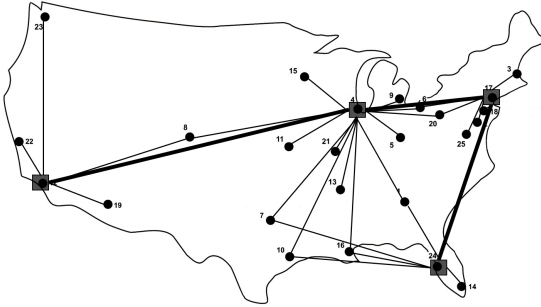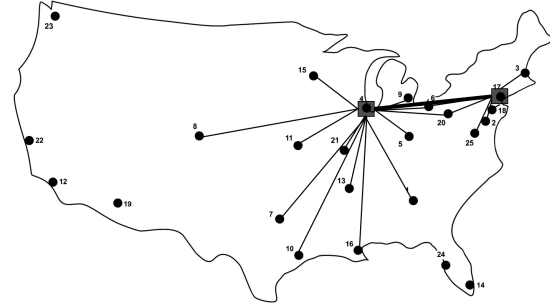
17

optimal solution of the problem.
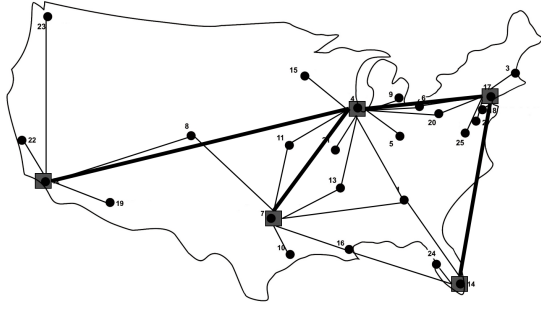
(a) High revenue, Low costs, $\alpha = 0.2$, Satisfied demand $= 100\%$
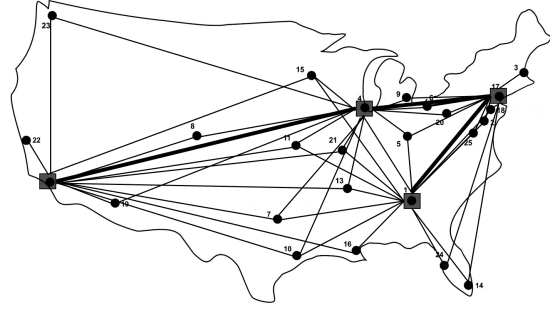
(b) High revenue, Low costs, $\alpha = 0.8$, Satisfied demand $= 86.83\%$

(c) High revenue, Medium costs, $\alpha = 0.2$, Satisfied demand $= 99.00\%$

(d) Low revenue, Medium costs, $\alpha = 0.2$, Satisfied demand $= 36.00\%$

Figure 1: Multiple allocation solutions.

Values of all the parameters except the α value are the same in Figures 1(a) and 1(b). Observe that the number of allocation links used in Figure 1(b) is much higher than that of Figure 1(a). This is because when $\alpha = 0.2$, most of the non-hub nodes are allocated to a single hub in order to take advantage from economies of scale, and, thus, decrease transportation cost of the flows via the inter-hub links. Whereas, when $\alpha = 0.8$, most of the non-hub nodes are allocated to at least two hubs. In this case, more flow is shipped using one hub on a route rather than using more hubs and inter-hub links. Moreover, when the discount due to economies of scale is high ($\alpha = 0.2$), the percentage of satisfied demand goes up to 100%.

Figure 1(a) corresponds to an instance with low cost level, whereas Figure 1(c) corresponds to medium costs. All the parameter values except the cost levels are the same in these two figures. When the cost level is increased from low to medium, note that one less hub is opened and the percentage of satisfied demand drops from 100% to 99%. The 1% of unsatisfied demand in Figure 1(c) refers to the demand between O-D pairs Dallas (7)–Seattle (23), Houston (10)–Seattle (23), and New Orleans (16)–Seattle (23). The demand between those city pairs are not served as it is not profitable.

The only difference between the instances depicted in Figures 1(c) and 1(d) is the revenue

level. When the revenue level is high (Figure 1c), four hubs are opened with 99% of total satisfied demand. On the other hand, when the revenue level is low (Figure 1d) the number of opened hubs reduces to two, and only 36% of the total demand is served. Note that some demand nodes in Figure 1(d) are not served at all. The demand generated from the cities on the East Coast and South–West (Los Angeles (12), Miami (14), Phoenix (19), San Francisco (22), Seattle (23), and Tampa (24)) remain entirely unserved in the optimal solution.

Figure 2 depicts instances from the single allocation solutions. In Figure 2(a), when the revenue level is high, five hubs are located and 99.68% of the demand is satisfied. The only unserved O-D pair is Denver(8)–Seattle (23) in this instance. On the other hand, in Figure 2(b), when the revenue level is low, one less hub is opened and the percentage of satisfied demand drops down to 56.33%. Similarly, when Figures 2(c) and 2(d) are compared, one less hub is opened and the percentage of satisfied demand decreases from 92.00% to 58.67% when the revenue level is reduced from high to medium.

(a) High revenue, Low costs, $\alpha = 0.2$, Satisfied demand = 99.67%

(b) Low revenue, Low costs, $\alpha = 0.2$, Satisfied demand = 57.33%

(c) High revenue, Medium costs, $\alpha = 0.4$, Satisfied demand = 92.00%

(d) Medium revenue, Medium costs, $\alpha = 0.4$, Satisfied demand = 58.67%

Figure 2: Single allocation solutions.
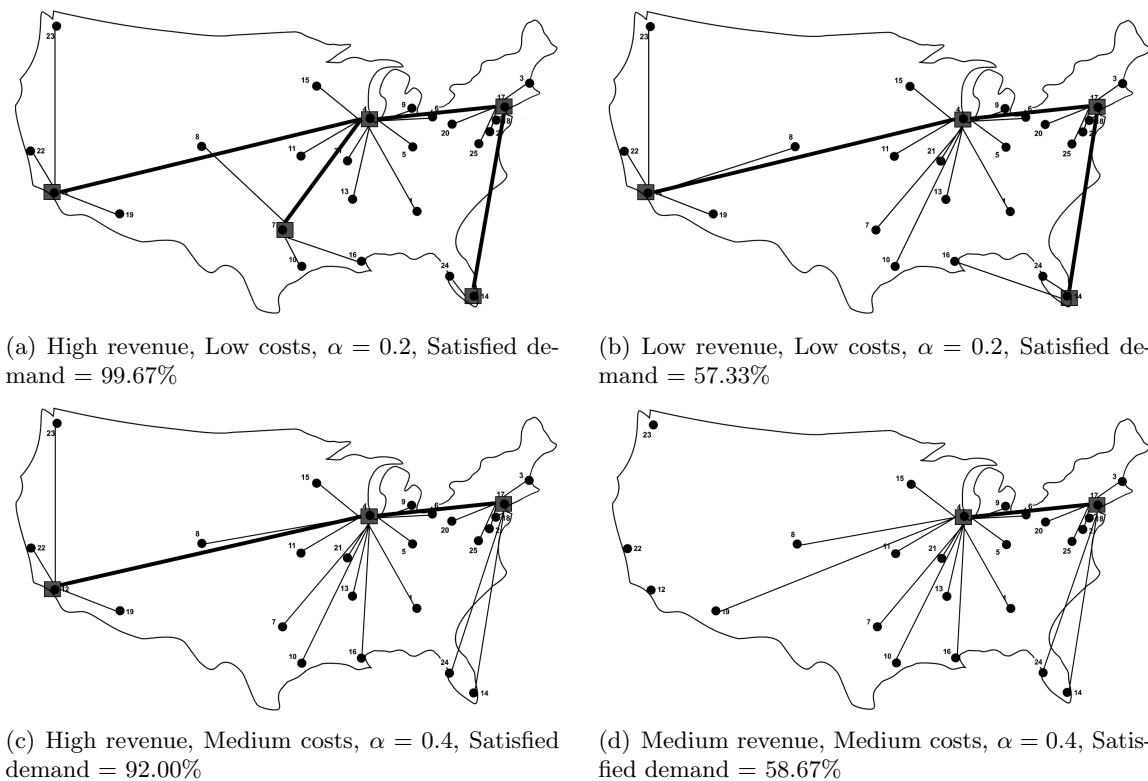
All the parameters in Figures 1(a) and 2(a) are the same; the only difference is the allocation strategies. Note that both of the models resulted in exactly the same hub locations and inter-hub networks. The differences between the resulting networks of these two instances are the consequence of the allocation strategies. In particular, Denver (8) is allocated to hubs located

in Dallas (7) and Los Angeles (12) with the multiple allocation model (Figure 1a), whereas it is allocated to a single hub located in Dallas (7) with the single allocation model (Figure 2a). All demand is satisfied in the multiple allocation solution, whereas the percentage of satisfied demand is 99.68% with single allocation. Allocating Denver (8) to Los Angeles (12) in addition to Dallas (7), makes it possible to serve the demand between the O-D pair Denver(8)–Seattle (23) in the optimal solution.

Observe from Figures 1 and 2 that Chicago (4), Los Angeles (12), and New York (17) are commonly preferred cities for locating hub nodes. Moreover, except for the 2-hub networks, note that the most profitable inter-hub networks are not fully interconnected, they are incomplete.

Next, we solve our models with direct connections under the same parameter values. The results for the multiple and single allocation models with direct connections are reported in Table 3. In addition to the columns in Table 2, columns labeled 'Satisfied demand–direct (%)" are included in Table 3 to highlight the percentage of the demand satisfied only through direct connections.

Similar conclusions can be drawn from Table 3 as from Table 2. An increase in the cost level results in a decrease in the percentage of satisfied demand and net profit. Besides, an increase in the α value, may also yield to a decrease in net profit. As expected, a decrease in the revenue level results in a decrease in net profit as well.

Multiple allocation with direct connections problem provides again an upper bound for the single allocation with direct connections problem. Thus, multiple allocation results in better net profits. Observe from Table 3 that in the instances with the same set of parameter values, percentage of satisfied demand with direct connections obtained with the single allocation model is higher than that of the multiple allocation model. That is, more demand is shipped through direct connections with the single allocation model compared with multiple allocation.

There are some instances with a hyphen sign (-) written under the hub locations columns in Table 3. No hubs are established in such instances, and, hence, the demand is satisfied only through direct connections. In this case, both allocation strategies yield the same net profit for the same set of parameter values.

Observe from Table 3 that percentage of satisfied demand through direct connections increases when fewer hubs are opened in the solutions. Moreover, when Tables 2 and 3 are compared, the models with direct connections result in opening fewer hubs. Fewer hubs are required with direct connections because a certain portion of the demand is served through direct service. Also note that the models with direct connections result in higher net profit values compared with the models not allowing for direct connections. If it was not profitable to serve the demand through direct connections, then no direct links would be established in

Table 3: Solutions allowing for direct connections with the CAB dataset.

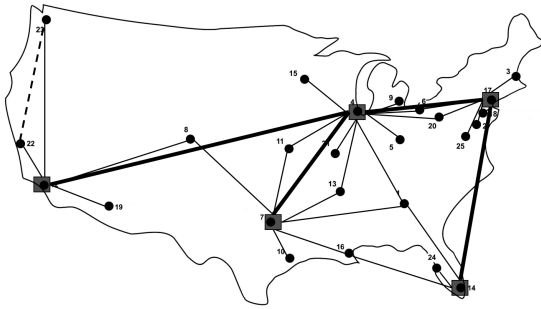| Costs	α	Multiple Allocation with Direct Connections					Single Allocation with Direct Connections				
		Net profit	Satisfied demand–total (%)	Satisfied demand–direct (%)	Hub locations	CPU time (s)	Net profit	Satisfied demand–total (%)	Satisfied demand–direct (%)	Hub locations	CPU time (s)
High revenue											
Low	0.2	1192	100.00	0.33	4,7,12,14,17	6537	1190	99.67	1.00	4,7,12,14,17	11628
	0.4	1042	97.83	0.67	4,7,12,14,17	6171	1027	95.33	3.00	4,12,14,17	9143
	0.6	945	94.50	2.33	4,7,12,25	2948	909	79.00	9.67	4,20	8457
	0.8	904	82.83	8.33	20,21	619	889	77.67	16.00	5	2018
Medium	0.2	925	99.00	1.33	4,12,18,24	3331	917	96.00	1.67	4,12,17	8675
	0.4	826	93.33	2.00	4,12,18	2666	814	91.67	2.33	4,12,18	7299
	0.6	779	76.00	7.00	5	264	779	76.00	7.00	5	1856
	0.8	779	76.00	7.00	5	27	779	76.00	7.00	5	1325
High	0.2	750	97.33	1.33	4,12,17	2461	738	96.67	1.33	4,12	5448
	0.4	696	69.33	4.33	20	88	696	69.33	4.33	20	1933
	0.6	696	69.33	4.33	20	35	696	69.33	4.33	20	1332
	0.8	696	69.33	4.33	20	41	696	69.33	4.33	20	1379
Medium revenue											
Low	0.2	694	94.67	0.67	4,7,12,14,17	4322	689	93.67	1.00	4,7,12,14,17	10296
	0.4	558	83.17	1.67	4,12,14,17	2868	547	81.33	3.33	4,12,14,17	8612
	0.6	505	67.00	5.33	4,18	536	497	59.00	11.33	20	1923
	0.8	497	59.00	11.33	20	159	497	59.00	11.33	20	1118
Medium	0.2	451	84.33	1.67	4,12,17	2491	431	83.00	1.67	4,12,17	4404
	0.4	401	55.67	6.00	20	223	401	55.67	6.00	20	1421
	0.6	401	55.67	6.00	20	44	401	55.67	6.00	20	1077
	0.8	401	55.67	6.00	20	36	401	55.67	6.00	20	957
High	0.2	325	54.33	3.33	20	34	325	54.33	3.33	20	1123
	0.4	325	54.33	3.33	20	31	325	54.33	3.33	20	891
	0.6	325	54.33	3.33	20	32	325	54.33	3.33	20	948
	0.8	325	54.33	3.33	20	31	325	54.33	3.33	20	919
Low revenue											
Low	0.2	213	49.33	2.00	4,14,17	2101	210	48.00	1.67	4,14,17	2527
	0.4	181	31.67	5.67	20	56	181	31.67	5.67	20	890
	0.6	181	31.67	5.67	20	42	181	31.67	5.67	20	783
	0.8	181	31.67	5.67	20	39	181	31.67	5.67	20	795
Medium	0.2	119	5.67	5.67	-	32	119	5.67	5.67	-	665
	0.4	119	5.67	5.67	-	29	119	5.67	5.67	-	618
	0.6	119	5.67	5.67	-	35	119	5.67	5.67	-	664
	0.8	119	5.67	5.67	-	36	119	5.67	5.67	-	635
High	0.2	89	4.33	4.33	-	32	89	4.33	4.33	-	699
	0.4	89	4.33	4.33	-	29	89	4.33	4.33	-	651
	0.6	89	4.33	4.33	-	31	89	4.33	4.33	-	654
	0.8	89	4.33	4.33	-	30	89	4.33	4.33	-	637

the resulting networks. Hence, allowing for direct connections provide an upper bound for no-direct-connection problems.

The single allocation with direct connections model is computationally more difficult than the multiple allocation version. The increase in run times is a consequence of the additional binary variable defined for the single allocation version of the problem. The average CPU time requirements for the instances listed in Table 3 is 17.81 minutes for the multiple allocation, and 48.33 minutes for the single allocation model.

Figure 3 provides optimal networks of some instances with direct connections when $\alpha = 0.2$. Figures 3(a) and 3(b) depict multiple allocation instances, while Figures 3(c) and 3(d)

depict their single allocation counterparts. The dash-lines in these figures indicate the direct connections between non-hub nodes. Note that resulting optimal inter-hub networks are all incomplete.

When the cost level is increased from low (Figures 3a and 3c) to medium (Figures 3b and 3d), fewer hubs are opened, and more demand is satisfied through direct connections. Figures 3(a) and 3(c) that correspond to multiple and single allocation problems, respectively, resulted in locating exactly the same hubs and the same inter-hub network design. However, the allocation links and the direct connections are different. In Figures 3(b) and 3(d), when the allocation strategy is changed from multiple to single, one less hub is opened and the percentage of satisfied demand through direct connections increased from 1.33% to 1.67%.

(a) Multiple allocation: High revenue, Low costs, Satisfied demand = 100% (0.33% with direct connections)
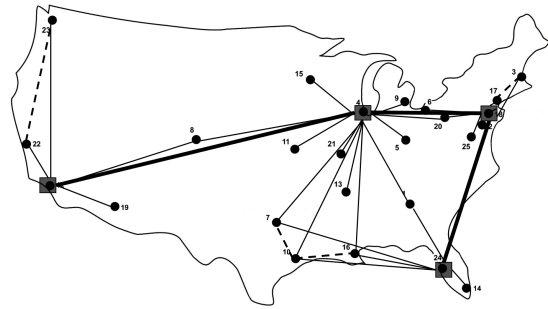
(b) Multiple allocation: High revenue, Medium costs, Satisfied demand = 99.00% (1.33% with direct connections)
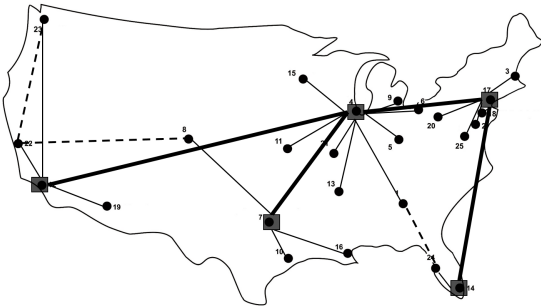
(c) Single allocation: High revenue, Low costs, Satisfied demand = 99.67% (1.00% with direct connections)

(d) Single allocation: High revenue, Medium costs, Satisfied demand = 96.00% (1.67% with direct connections)

Figure 3: Multiple and single allocation solutions with direct connections when $\alpha = 0.2$.

Next, we compare instances presented in Figures 1, 2, and 3. Note that the values of all the parameters are the same in Figures 1(a), 2(a), 3(a), and 3(c). Each of these figures corresponds to an optimal solution obtained under a different model. However, all optimal solutions resulted in opening the same set of hubs and inter-hub links. The only differences are the allocation links and direct connection decisions. Shipping the demand between San Francisco (22) and

Seattle (23) directly as seen in Figure 3(a) resulted in an increase in the net profit value from 1190 to 1192 compared with Figure 1(a). Total percentage of satisfied demand in the instances depicted in Figures 2(a) and 3(c) are exactly the same (99.67%). However, 1% of the demand is satisfied through direct connections in Figure 3(c), and this resulted in an increase in the net profit from 1181 to 1190.

Table 4: r-allocation solutions with the CAB dataset.

Cost	α	r-Allocation without Direct Connections				r-Allocation with Direct Connections			
		$r=2$		$r=3$		$r=2$		$r=3$	
		Net profit	CPU time (s)	Net profit	CPU time (s)	Net profit	CPU time (s)	Net profit	CPU time (s)
					High revenue				
Low	0.2	1187	20788	1190	21895	1191	20001	1192	21174
	0.4	1021	20700	1033	21134	1034	19108	1042	20602
	0.6	897	18534	926	20409	921	16800	945	17418
	0.8	835	15671	850	17381	904	4149	904	4287
Medium	0.2	905	18890	913	19924	925	18144	925	19147
	0.4	804	16423	804	16562	826	16722	826	17015
	0.6	735	14544	735	15276	779	1312	779	1245
	0.8	681	10500	681	10662	779	1216	779	1422
High	0.2	740	14719	740	15283	750	14603	750	15067
	0.4	634	11466	634	12084	696	1301	696	1385
	0.6	599	2140	599	2266	696	1434	696	1283
	0.8	599	2130	599	2160	696	1430	696	1323
					Medium revenue				
Low	0.2	687	17332	690	18037	692	19527	694	21001
	0.4	542	16749	550	17712	552	16899	558	19084
	0.6	456	14959	456	15134	505	4897	505	5242
	0.8	426	14279	426	14623	497	1295	497	1448
Medium	0.2	429	7578	429	7982	451	13212	451	14154
	0.4	349	7053	349	7379	401	1898	401	2138
	0.6	328	4741	328	5044	401	2012	401	1980
	0.8	325	3191	325	3335	401	1827	401	2001
High	0.2	267	5763	267	6182	325	1416	325	1461
	0.4	260	1786	260	1644	325	1323	325	1526
	0.6	260	1500	260	1587	325	1284	325	1315
	0.8	260	1401	260	1509	325	1217	325	1542
					Low revenue				
Low	0.2	195	12633	199	14689	213	4010	213	4142
	0.4	157	1913	157	2133	181	779	181	785
	0.6	142	1332	142	1428	181	764	181	791
	0.8	132	804	132	1044	181	873	181	803
Medium	0.2	69	825	69	998	119	623	119	675
	0.4	65	205	65	225	119	651	119	602
	0.6	65	187	65	243	119	604	119	634
	0.8	65	186	65	214	119	652	119	704
High	0.2	15	45	15	49	89	623	89	637
	0.4	15	43	15	46	89	640	89	662
	0.6	15	43	15	44	89	638	89	651
	0.8	15	42	15	46	89	681	89	676

We also want to analyze the performance of the r-allocation models. Looking at the optimal multiple allocation solutions, we observed that a non-hub node is allocated to at most three

hubs with the CAB dataset. Therefore, when $r = 3$, r-allocation problems reduce to multiple allocation on this dataset. We thus set r to 2 and 3 to take runs with the two r-allocation models under different parameter values. Table 4 presents the results obtained with the r-allocation models.

Note that, when $r = 3$, r-allocation models result in exactly the same optimal solutions obtained with the multiple allocation models. Moreover, when $r = 2$, most of the instances again resulted in the same optimal solutions with $r = 3$. Hence, we only report net profit values and CPU times in Table 4. Note that instances with $r = 3$ provides an upper bound for $r = 2$. Thus, for the same set of parameters, when r value is increased from 2 to 3, the net profit value either increases or remains the same.

Regarding the CPU time requirements, r-allocation models turned out to be the most challenging set of models compared with single and multiple allocation. They are computationally more difficult than multiple allocation models because an additional set of binary variables needs to be defined. Compared with single allocation, allowing the models search for r allocations instead of just one seems to increase the CPU times considerably as well. The most time-consuming instance with the r-allocation models took a little less than 6 hours to solve to optimality. The averages of the CPU times listed in Table 4 is 1.89 hours.
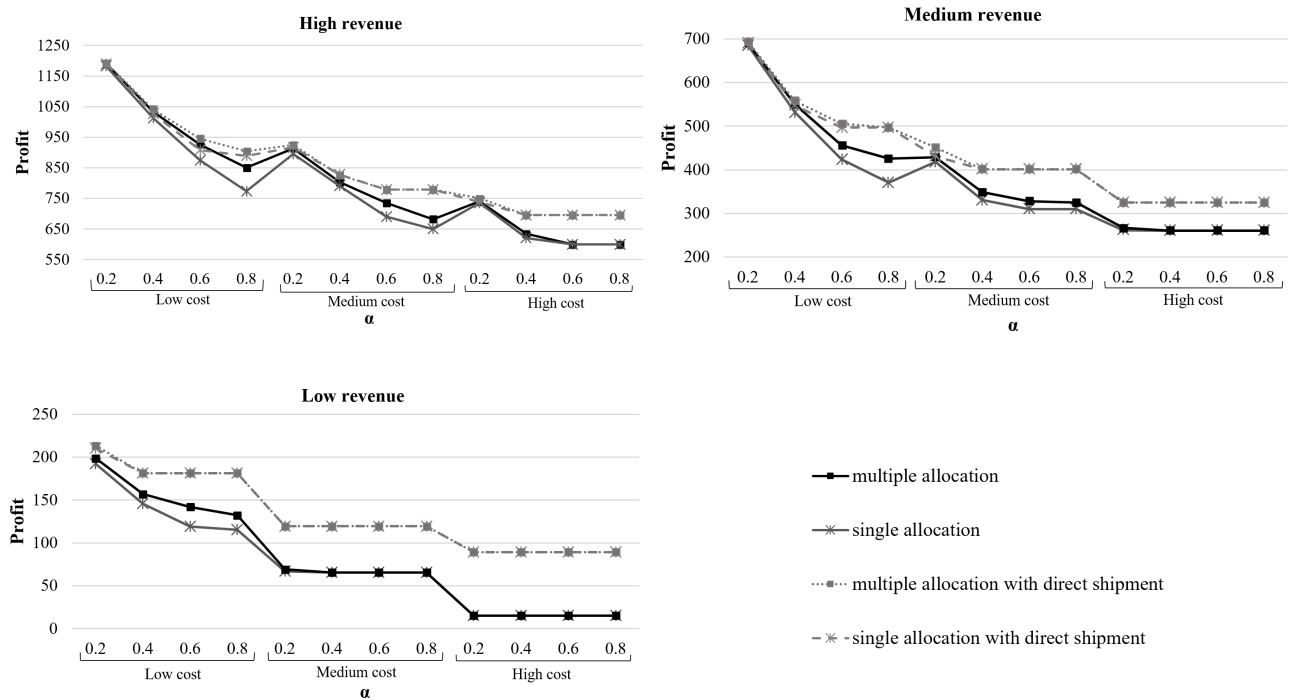

Figure 4: Profit comparison under different parameter values with four different models.

Net profits change significantly with the cost levels and the economies of scale factor under

different allocation strategies. Figure 4 provides an insight for the change in net profits under different parameter values.

Each graph presented in Figure 4 corresponds to a different level of revenue: high, medium, and low. For each revenue level, we depict optimal net profit values obtained with four different models under different cost levels and discount factor α.

Observe from Figure 4 that multiple allocation model with direct connections results in the best net profit values. As already discussed, this model actually provides an upper bound for all the remaining models. The lowest profits are obtained with the single allocation model when direct connections are not allowed.

Figure 4 also illustrates the impact of α on net profit. It is clear that profits decrease with increasing α, for a given cost level. The effect of α on profits is higher with lower cost levels. Note that there are some instances in which the profits are insensitive to α. In such instances, either no hub or one hub is opened in the optimal solutions. In other words, no economies of scale is achieved as there is no inter-hub transportation. In such cases, multiple and single allocation strategies result in exactly the same net profits.

In general, we can conclude from Figure 4 that when the effect of economies of scale on transportation costs is higher (when α value is lower), the decision maker can obtain significantly more profit. Hence, economies of scale is an important factor in designing and operating hub networks. As noted in the problem definition, in this study, we assume that economies of scale is exploited only on the inter-hub links and that it is independent from the amount of flow. This is a simplification of real-life for the sake of modeling the problem efficiently. We next want to analyze the consequences of this simplification.

5.2   Economies of scale analysis

We analyzed optimal flows on all network connections in the resulting solutions. The amount of flow routed on the inter-hub links and allocation connections are obviously less when direct connections between non-hub nodes are allowed. Hence, we concentrated on the optimal flows with the models allowing for direct connections to observe if the amount of flow routed on the inter-hub links justify economies of scale. For this analysis, we sorted flows on all links of the network to see if there are any allocation links or direct links carrying larger flows than inter-hub links. We could identify only a few such instances where the flow on one allocation link exceeded the flow on an inter-hub link. For example, in the instance from Table 3 with high revenue and medium cost levels with $\alpha = 0.4$, the allocation link New York (17)–Philadelphia (18) carries more flow than the inter-hub link Chicago (4)–Los Angeles (12). This is because all demand originating at New York and destined to all other cities has to use the allocation link from

New York to Philadelphia. However, as mentioned above, only a few instances do not justify economies of scale in our experimentation. This is because our models do not enforce the establishment of fully interconnected inter-hub networks. Note that in all of our solutions with more than two hubs, the resulting inter-hub networks are incomplete. By the inclusion of inter-hub network design decisions in hub location models, more flow is consolidated on hub-to-hub links.

We also wanted to analyze the impact of using a flow-independent economies of scale factor on the inter-hub links. Incorporating a flow-dependent economies of scale factor in our models would computationally be impractical due to its non-linear nature. Instead, we decided to recalculate the transportation costs of our optimal solutions by using flow-dependent discounts. For this analysis, we adopted the non-linear concave function introduced in O'Kelly and Bryan [48]. While recalculating the transportation costs, we replaced α and for each inter-hub link $k, l \in N$, we calculated the discount through the function:

$$1 - \theta(\sum_{i \in N} f_{ikl})^{\beta} \qquad k, l \in N \tag{34}$$

where $0 < \theta \leq 1$, $\beta > 0$, and $\sum_{i \in N} f_{ikl}$ is the amount of flow routed on the inter-hub link from hub k to hub l. Note that since total demand is scaled and it adds up to one, we do not need to scale the flows in this calculation.

For each instance with the given revenue and cost levels, we compared net profits obtained from using different combinations of constant and flow-dependent discount factors, α, θ, and β. We observed through this analysis that for any given value of α, there is a valid combination of θ and β values which results in comparable transportation costs. For example, in the instances with high revenue and low cost levels, total transportation cost recalculated by using $\theta = 0.9$ and $\beta = 0.1$ in the above function results in a 2% decrease in the net profit on average when compared with using $\alpha = 0.2$, which is the highest economies of scale discount that we used in our experimentation. Similarly, with high revenue and medium cost levels, using the combination of $\theta = 0.75$ and $\beta = 0.2$ results in only a 0.23% decrease in the net profit on average compared with using $\alpha = 0.4$. This analysis shows that the α values used in our computational experiments actually provide good estimates for flow-dependent discounts.

We would like to note that the analysis that we performed with this flow-dependent discount function is post-processing. As noted by O'Kelly and Bryan [48], modeling the problem assuming a fixed discount factor not only miscalculate the total network cost but may also wrongly select optimal hub locations and allocations. Incorporating a more realistic calculation of economies of scale within the proposed models is definitely an important avenue for future research.

5.3 Comparison with the literature

As discussed in our introduction, hub location problems with profits considering the multiple allocation strategy was already introduced in Alibeyg et al. [1]. In this paper, in addition to studying all allocation possibilities, we included additional design variables to account for the design of the inter-hub network which in turn makes the problem more complicated. To justify the increased complexity of the proposed models, we present a computational comparison of our multiple allocation model and the PO_1 model presented in Alibeyg et al. [1, 2]. For this comparison, we concentrated on the instances resulted in opening more than two hubs in the optimal solutions obtained from our multiple allocation model. Table 5 presents the results.

Table 5: Comparison of our multiple allocation model with PO_1 from Alibeyg et al. [1, 2].

Cost	alpha	PO_1				Multiple Allocation Model				Profit
		Net profit	Satisfied demand (%)	LP gap (%)	CPU time (s)	Net profit	Satisfied demand (%)	LP gap (%)	CPU time (s)	gap (%)
					High revenue					
Low	0.2	1146	100.00	3.49	1851	1190	100.00	2.73	4891	3.86
	0.4	991	95.67	3.61	1611	1033	97.33	4.31	4130	4.23
	0.6	889	91.67	5.28	1084	926	94.00	2.77	2428	4.19
	0.8	832	85.33	7.95	449	850	86.83	2.27	754	2.21
Medium	0.2	887	97.05	6.05	764	913	99.00	7.31	801	2.96
	0.4	791	88.33	4.91	416	804	93.00	5.84	615	1.62
	0.6	723	85.67	9.09	287	735	89.33	3.70	521	1.63
High	0.2	723	93.33	6.22	325	740	97.00	8.04	550	2.27
	0.4	619	88.33	13.52	144	634	93.00	9.45	362	2.31
					Medium revenue					
Low	0.2	656	88.00	3.35	1081	690	94.00	4.91	1756	5.16
	0.4	517	83.00	6.95	605	550	89.00	6.43	1249	6.36
	0.6	440	64.00	4.24	427	456	68.00	3.49	504	3.65
	0.8	411	61.00	6.61	212	426	64.00	2.85	283	3.56
Medium	0.2	401	82.00	17.73	324	429	83.17	12.63	549	6.98
					Low revenue					
Low	0.2	183	55.33	8.99	111	199	58.83	4.79	80	8.74

For each instance, Table 5 provides the net profit, percentage of satisfied demand, LP relaxation gap, and the run time of the instances obtained by solving the PO_1 model from Alibeyg et al. [1, 2] and our multiple allocation model introduced in Section 4.1. The last column of Table 5 presents the percentage of the gap between the net profits obtained from these two models.

Firstly, a comparison of the CPU times and LP relaxation gaps reported in Alibeyg et al. [1] with Table 5 reveals that our test instances are computationally more challenging than the ones reported in Alibeyg et al. [1]. We would like to note that we used variable fixing while solving both of the models under the default settings of CPLEX.

Our multiple allocation model provides an upper bound in terms of the net profit value and

27

the percentage of satisfied demand. Observe from Table 5 that all the profit gaps are strictly positive. This is because in all of our solutions with more than two hubs, there exist paths in which more than two hubs are used to satisfy the demand between O-D pairs. PO_1 model needs to activate additional hub links to satisfy the same or even less amount of demand that is satisfied with our multiple allocation model. For example, with high revenue and low cost levels when $\alpha = 0.2$, 100% of the demand is satisfied under both of the models. In this instance, our multiple allocation solution resulted in operating an incomplete inter-hub network (Figure 1a) and 14% of the total flow is routed using more than two hubs on a route. While, the PO_1 solution, which restricts the number of hubs on a route to at most two and activates hub arcs of a fully inter-connected hub network, resulted in a 3.83% decrease in the net profit. The percent gap in profits goes up to 8.74% in the instances listed in Table 5. Moreover, since the demand is scaled in our data, the magnitude of the objective function value is relatively low in our experiments, otherwise the difference in the net profits would have been higher. We believe these results provide a clear indication of the added benefit of incorporating more complex paths into the models.

5.4   Results with the AP dataset

In this section, we test the performance of our models on a 40-node subset of the Australia Post (AP) dataset (Ernst and Krishnamoorthy [29]). The distances and the demand between each pair of cities are provided in OR Library (Beasley [12]). Collection and distributions costs per unit are taken equal to one. For the revenue of each O-D pair, we again test three different values such that $r_{ij} \in \{20, 30, 50\}$ for all $i, j \in N$. We refer to these values as low, medium, and high revenues, respectively. There are two different sets for hub installation costs available on the AP dataset referred to as loose and tight. Cost of operating an inter-hub link is assumed to be 10% of the average installation costs from the AP data. Similar to the CAB dataset, the cost of operating a direct link is taken to be the 20% of the operational cost of an inter-hub link. Operational costs are assumed to be the same for all potential links. For all instances, the discount factor α is taken as 0.75, as defined in the AP dataset (Beasley [12]).

The results obtained from solving the multiple and single allocation models without and with direct connections are summarized in Tables 6 and 7, respectively. As noted in the first two columns, we tested three different revenue and two different cost levels.

Similar conclusions can be drawn with the AP dataset as with the CAB dataset. Net profits and the percentage of satisfied demand decrease with decreasing revenue levels. Tight (high) fixed cost instances result in lower net profits. In the models with direct connections (Table 7), a decrease in the revenue level results in an increase in the percentage of satisfied demand through

Table 6: Multiple and single allocation solutions with the AP dataset.

Revenue	Costs	Multiple Allocation				Single Allocation			
		Net profit	Satisfied demand (%)	Hub locations	CPU time (h)	Net profit	Satisfied demand (%)	Hub locations	CPU time (h)
High	Loose	118303	98.13	12,22,26,28	3.82	73384	91.25	11,12,22,28	6.87
High	Tight	105850	97.75	11,14,29	3.07	46736	90.38	11,14,29	6.12
Medium	Loose	42868	77.69	12,22,26,28	2.91	36904	64.06	11,22,28	5.04
Medium	Tight	32716	69.13	14,29	0.09	29569	51.56	29	1.20
Low	Loose	14099	22.63	28	0.03	14099	22.63	28	0.53
Low	Tight	6404	23.50	29	0.05	6404	23.50	29	0.48

Table 7: Solutions allowing for direct connections with the AP dataset.

Revenue	Costs	Multiple Allocation with Direct Connections					Single Allocation with Direct Connections				
		Net profit	Satisfied demand-total (%)	Satisfied demand-direct (%)	Hub locations	CPU time (h)	Net profit	Satisfied demand-total (%)	Satisfied demand-direct (%)	Hub locations	CPU time (h)
High	Loose	119998	98.50	1.00	12,18,22,28	4.61	76026	87.00	4.75	11,12,22,28	7.01
High	Tight	106614	95.25	1.13	14,29	3.66	48101	86.13	5.38	11,14,29	6.59
Medium	Loose	44222	76.06	1.19	11,22,28	3.05	38694	63.19	5.56	11,22,28	5.82
Medium	Tight	36252	71.00	5.25	19,22	0.57	34179	54.75	12.31	29	2.15
Low	Loose	15077	23.44	1.31	28	0.04	15077	23.44	1.31	28	0.39
Low	Tight	6677	23.63	0.25	29	0.09	6677	23.63	0.25	29	0.34

direct connections.

The CPU times obtained from solving the models with the AP dataset indicate that larger-sized instances are more challenging, and, consequently, more time consuming than small-sized instances, as expected. Because of the higher run times, CPU times in Tables 6 and 7 are reported in hours. Note that, in addition to the size of the instances, computation times also vary a lot with different parameter values as well as allocation strategies. In particular, when the revenue level decreases from high to low, the CPU times in Table 6 drops from about 7 hours to less than 1 hour. The most difficult instances were with the single allocation model allowing for direct connections. Average CPU time requirement of the instances presented in Tables 6 and 7 with the 40-node AP dataset is 2.66 hours.

6 Conclusions

We studied hub location problems with the aim of maximizing profit. We defined new problems and developed efficient mathematical models addressing the questions of where to locate hubs and how to design hub networks in order to maximize profit. We determined optimal routes of

flow to transport the demand between origin-destination pairs while allowing a portion of the demand to be unserved. We introduced mixed-integer linear programming models for multiple, single, and r-allocation versions of the problem. We also modeled the possibility of direct connections for each allocation strategy. The proposed models design hub networks in the most profitable way considering all possibilities for shipments.

We performed computational experiments on the well-known CAB and AP datasets to analyze the resulting hub networks and performance of the proposed mathematical models. We tested all models under various different parameter settings. We varied revenues, costs, and the economies of scale discount factor, and analyzed the optimal locations of hubs, design of the hub network, and percentage of satisfied demand.

The results provided insights on where to locate hubs, how to design hub networks, what portion of the demand to serve, and how to route flows. The most profitable inter-hub network designs with more than two hubs turned out to be incomplete. Trade-off between different allocation strategies as well as the impact of allowing for direct connections is explored. The best net profit values were obtained with the multiple allocation model allowing for direct connections, whereas, the lowest profits were obtained with the single allocation model when direct connections are not allowed. The choice of the allocation strategy did not result in a significant difference on the locations of hubs in the optimal solutions. The effect of the economies of scale discount on total profit is also analyzed. The results showed that the decision maker can obtain significantly more profit when the discount on transportation costs due to economies of scale is higher.

For future research, exact solution methodologies or heuristic algorithms can be developed for the introduced problems. Given the computational effort required to solve some of the instances, efficient solution algorithms are needed to solve realistically-sized instances. Another future research direction would be to incorporate a more realistic calculation of economies of scale into the proposed models. Embedding uncertainty in the models is another area of research which is definitely worth pursuing.

Acknowledgment

References

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

[2] Alibeyg, A., Contreras, I., Fernández, E., 2017. Exact solution of hub network design problems with profits. European Journal of Operational Research 000, 1–15.

[3] Alumur, S., Kara, B. Y., 2008. Network hub location problems: The state of the art. European Journal of Operational Research 190 (1), 1–21.

[4] Alumur, S., Yaman, H., Kara, B., 2012. Hierarchical multimodal hub location problem with time-definite deliveries. Transportation Research Part E 48, 1107–1120.

[5] Alumur, S. A., Kara, B. Y., Karasan, O. E., 2009. The design of single allocation incomplete hub networks. Transportation Research Part B: Methodological 43 (10), 936–951.

[6] Alumur, S. A., Kara, B. Y., Karasan, O. E., 2012. Multimodal hub location and hub network design. Omega 40 (6), 927–939.

[7] Alumur, S. A., Nickel, S., Saldanha-da Gama, F., 2012. Hub location under uncertainty. Transportation Research Part B: Methodological 46 (4), 529–543.

[8] Alumur, S. A., Nickel, S., Saldanha-da Gama, F., Seçerdin, Y., 2016. Multi-period hub network design problems with modular capacities. Annals of Operations Research 246 (1-2), 289–312.

[9] Aykin, T., 1994. Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem. European Journal of Operational Research 79 (3), 501–523.

[10] Aykin, T., 1995. Networking policies for hub-and-spoke systems with application to the air transportation system. Transportation Science 29 (3), 201–221.

[11] Azizi, N., Chauhan, S., Salhi, S., Vidyarthi, N., 2016. The impact of hub failure in hub-and-spoke networks: Mathematical formulations and solution techniques. Computers & Operations Research 65, 174–188.

[12] Beasley, J. E., 1990. OR Library: Hub location. http://people.brunel.ac.uk/mastjjb/jeb/ orlib/phubinfo.html.

[13] Boland, N., Krishnamoorthy, M., Ernst, A. T., Ebery, J., 2004. Preprocessing and cutting for multiple allocation hub location problems. European Journal of Operational Research 155 (3), 638–653.

[14] Campbell, J. F., 1994. Integer programming formulations of discrete hub location problems. European Journal of Operational Research 72 (2), 387–405.

[15] Campbell, J. F., 2013. Modeling economies of scale in transportation hub networks. 46th Hawaii International Conference on System Sciences, 1154–1163.

[16] Campbell, J. F., Ernst, A. T., Krishnamoorthy, M., 2002. Hub location problems. Facility Location: Applications and Theory, 373–407.

[17] Campbell, J. F., Ernst, A. T., Krishnamoorthy, M., 2005. Hub arc location problems: Part I - Introduction and Results. Management Science 51 (10), 1540–1555.

[18] Campbell, J. F., Ernst, A. T., Krishnamoorthy, M., 2005b. Hub arc location problems: Part II - Formulations and Optimal Algorithms. Management Science 51 (10), 1556–1571.

[19] Campbell, J. F., O'Kelly, M. E., 2012. Twenty-five years of hub location research. Transportation Science 46 (2), 153–169.

[20] Chen, J.-F., 2007. A hybrid heuristic for the uncapacitated single allocation hub location problem. Omega 35 (2), 211–220.

[21] Contreras, I., 2015. Hub location problems. In: Location Science. Springer, pp. 311–344.

[22] Contreras, I., Cordeau, J.-F., Laporte, G., 2011. The dynamic uncapacitated hub location problem. Transportation Science 45 (1), 18–32.

[23] Contreras, I., Cordeau, J.-F., Laporte, G., 2011. Stochastic uncapacitated hub location. European Journal of Operational Research 212, 518–528.

[24] Contreras, I., Cordeau, J.-F., Laporte, G., 2011a. Benders decomposition for large-scale uncapacitated hub location. Operations Research 59 (6), 1477–1490.

[25] Contreras, I., Fernández, E., Marín, A., 2010. The tree of hubs location problem. European Journal of Operational Research 202 (2), 390–400.

[26] Contreras, I., Tanash, M., Vidyarthi, N., 2013. The cycle hub location problem. Technical Report CIRRELT 59.

[27] Correia, I., Nickel, S., Saldanha-da Gama, F., 2017. A stochastic multi-period capacitated multiple allocation hub location problem: Formulation and inequalities. Omega 74, 122–134.

[28] Eiselt, H. A., Marianov, V., 2009. A conditional p-hub location problem with attraction functions. Computers & Operations Research 36 (12), 3128–3135.

[29] Ernst, A. T., Krishnamoorthy, M., 1996. Efficient algorithms for the uncapacitated single allocation p-hub median problem. Location Science 4 (3), 139–154.

[30] Ernst, A. T., Krishnamoorthy, M., 1998a. Exact and heuristic algorithms for the uncapacitated multiple allocation p-hub median problem. European Journal of Operational Research 104 (1), 100–112.

[31] Gelareh, S., Nickel, S., Pisinger, D., 2010. Liner shipping hub network design in a competitive environment. Transportation Research Part E: Logistics and Transportation Review 46 (6), 991–1004.

[32] Horner, M. W., O'Kelly, M. E., 2001. Embedding economies of scale concepts for hub network design. Journal of Transport Geography 9 (4), 255–265.

[33] Hwang, Y. H., Lee, Y. H., 2012. Uncapacitated single allocation p-hub maximal covering problem. Computers & Industrial Engineering 63 (2), 382–389.

[34] Kim, H., O'Kelly, M. E., 2009. Reliable p-hub location problems in telecommunication networks. Geographical Analysis 41 (3), 283–306.

[35] Kimms, A., 2006. Economies of scale in hub & spoke network design models: We have it all wrong. In: Perspectives on Operations Research. Springer, pp. 293–317.

[36] Labbé, M., Yaman, H., Gourdin, E., 2005. A branch and cut algorithm for hub location problems with single assignment. Mathematical programming 102 (2), 371–405.

[37] Lüer-Villagra, A., Marianov, V., 2013. A competitive hub location and pricing problem. European Journal of Operational Research 231 (3), 734–744.

[38] Mahmutoğulları, A. İ., Kara, B. Y., 2015. Hub location problem with allowed routing between nonhub nodes. Geographical Analysis 47 (4), 410–430.

[39] Marianov, V., Serra, D., ReVelle, C., 1999. Location of hubs in a competitive environment. European Journal of Operational Research 114 (2), 363–371.

[40] Marin, A., Cánovas, L., Landete, M., 2006. New formulations for the uncapacitated multiple allocation hub location problem. European Journal of Operational Research 172 (1), 274–292.

[41] Martins de Sá, E., Camargo, R. S., Miranda, G., 2013. An improved Benders decomposition algorithm for the tree of hubs location problem. European Journal of Operational Research 226 (2), 185–202.

[42] Martins de Sá, E., Contreras, I., Cordeau, J.-F., Saraiva de Camargo, R., de Miranda, G., 2015. The hub line location problem. Transportation Science 49 (3), 500–518.

[43] Meraklı, M., Yaman, H., 2016. Robust intermodal hub location under polyhedral demand uncertainty. Transportation Research Part B: Methodological 86, 66–85.

[44] Nickel, S., Schöbel, A., Sonneborn, T., 2001. Hub location problems in urban traffic networks. In: Mathematical Methods on Optimization in Transportation Systems. Springer, pp. 95–107.

[45] O'Kelly, M. E., 1986. The location of interacting hub facilities. Transportation Science 20 (2), 92–106.

[46] O'Kelly, M. E., 1987. A quadratic integer program for the location of interacting hub facilities. European Journal of Operational Research 32 (3), 393–404.

[47] O'Kelly, M. E., 1992. Hub facility location with fixed costs. Papers in Regional Science 71 (3), 293–306.

[48] O'Kelly, M. E., Bryan, D., 1998. Hub location with flow economies of scale. Transportation Research Part B: Methodological 32 (8), 605–616.

[49] O'Kelly, M. E., Bryan, D., Skorin-Kapov, D., Skorin-Kapov, J., 1996. Hub network design with single and multiple allocation: A computational study. Location Science 4 (3), 125–138.

[50] O'Kelly, M. E., Campbell, J. F., Camargo, R. S., Miranda, G., 2015. Multiple allocation hub location model with fixed arc costs. Geographical Analysis 47 (1), 73–96.

[51] Peker, M., Kara, B. Y., 2015. The p-hub maximal covering problem and extensions for gradual decay functions. Omega 54, 158–172.

[52] Podnar, H., Skorin-Kapov, J., Skorin-Kapov, D., 2002. Network cost minimization using threshold-based discounting. European Journal of Operational Research 137 (2), 371–386.

[53] Sasaki, M., Campbell, J. F., Krishnamoorthy, M., Ernst, A. T., 2014. A stackelberg hub arc location model for a competitive environment. Computers & Operations Research 47, 27–41.

[54] Sasaki, M., Fukushima, M., 2001. Stackelberg hub location problem. Journal of the Operations Research Society of Japan 44 (4), 390–402.

[55] Serper, E. Z., Alumur, S. A., 2016. The design of capacitated intermodal hub networks with different vehicle types. Transportation Research Part B: Methodological 86, 51–65.

[56] Skorin-Kapov, D., Skorin-Kapov, J., O'Kelly, M., 1996. Tight linear programming relaxations of uncapacitated p-hub median problems. European Journal of Operational Research 94 (3), 582–593.

[57] Sung, C., Jin, H., 2001. Dual-based approach for a hub network design problem under non-restrictive policy. European Journal of Operational Research 132 (1), 88–105.

[58] Tran, T. H., OHanley, J. R., Scaparra, M. P., 2016. Reliable hub network design: Formulation and solution techniques. Transportation Science 51 (1), 358–375.

[59] Wagner, B., 2007. An exact solution procedure for a cluster hub location problem. European Journal of Operational Research 178 (2), 391–401.

[60] Yaman, H., 2009. The hierarchical hub median problem with single assignment. Transportation Research Part B: Methodological 43 (6), 643–658.

[61] Yaman, H., 2011. Allocation strategies in hub networks. European Journal of Operational Research 211 (3), 442–451.

[62] Yaman, H., Elloumi, S., 2012. Star p-hub center problem and star p-hub median problem with bounded path lengths. Computers & Operations Research 39 (11), 2725–2732.