E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. x, No. x (xx) x–x.

Paper Type: Original Article

A variable service rate queue model for hub median problem

Sarow Saeedi[1]; Omid Poursabzi[2]; Zaniar Ardalan[3]; Sajad Karimi[4]

1 Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran
2 Department of Industrial Engineering, Faculty of Engineering, Kharazmi University
3 Systems and Industrial Engineering Department, University of Arizona, Tucson, Arizona, USA
4 Department of System Science and Industrial Engineering, Binghamton University, Binghamton, New York, USA

Citation:

LastName, (Abbreviation of FirstName), & LastName, (Abbreviation of FirstName). (2021). A variable service rate queue model for hub median problem. *Journal of applied research on industrial engineering, Volume* (Issue), PP.

Received: Reviewed: Revised: Accept:

Abstract

Hub location problems (HLP) have multiple applications in logistic systems, the airways industry, supply chain network design, and telecommunication. In the HLP, the selected nodes as hubs perform the principal role in processing the inflow arising from other nodes. So, congestion would be inevitable at hub nodes. This paper considers a p-hub median problem with multiple hub node servers delivering service at variable rates. Since the service rates are limited and variable, a queue is formed at each hub server. To tackle this problem, we developed a mixed-integer linear programming model that optimizes the selected hub nodes to reduce congestion under an allowable defined queue length at each server and minimize the total costs of the model, including transportation and hub establishment costs. We utilized the Civil Aeronautics Board (CAB) dataset containing 25 USA cities, which is a valuable source for designing numerical examples in the HLP, to prove the model's efficiency. The results obtained from the designed sample problems show that strategic decisions on defining the number of hubs and maximum acceptable queue length at each hub server will significantly impact the hub location network design.

Keywords: hub location problem, p-hub median, queuing system, congestion, variable service rate

Licensee **Journal of Applied Research on Industrial Engineering**. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0).

1 | Introduction

Hub location problem (HLP) involves designing a network consisting of multiple locations with inflows and outflows. The goal is to design a network to minimize the distribution and delivery costs while minimizing the costs of routes establishment between locations. To achieve this goal and design an optimal network of nodes, some locations will be chosen as hub locations, which act as the interface, with other locations connected to them. In other words, mutual transportation flow between nodes should be established.

Corresponding Author: Omid.Poursabzi@gmail.com

http://dx.doi.org/10.22105/jarie.2021.301634.1372

In Press

The established flow between any two nodes (hub or non-hub nodes) must be established through the hub nodes without using any intermediate node. On the other hand, each established flow should pass through at least one arc and, utmost, two arcs.

There exist four principal types of hub location problems in the literature, which are 1) capacitated and uncapacitated hub location problem, 2) p-hub median problem, 3) p-hub center problem, and 4) hub covering location problems.

The capacity of hub nodes may be capacitated/limited (LHLP) or uncapacitated/unlimited (UHLP). P-hub median problem (pHMP) minimizes the total cost of transportation in the network by optimally locating p hubs. In this case, the number of hubs is given. The main goal in the p-hub center problem (pHCP) is to minimize the farthest distance by finding the optimal location of p hubs and the allocation of non-hub nodes to the hubs. The number of hubs is unknown in the hub covering location problem (HCLP), and the main issues here are locating hubs and allocating non-hub nodes to them. This problem also includes cover constraints, meaning that each hub has a limitation on the number of non-hub nodes it can serve. Generally, hub location problems can be categorized into two types, i.e., single and multiple-allocations problems. In the former, each non-hub node should be allocated to precisely one hub, while in the latter, a non-hub node can be allocated to more than one hub.

According to the surveyed literature in the HLP, the p-hub median and the p-hub center location problems seem far more appealing to the researchers. However, the p-hub center location problems are more useful when studying emergency facility locations or circumstances concerning perishable products in transportation networks [1]. The primary goal in this paper is not to curb the scope of the study to cover merely such kinds of problems. Therefore, this paper addresses a p-hub median problem with hubs with a limited number of serves that process the network inflow at variable service rates. State-dependent service rates simulate the real-world conditions, as servers, in practical situations, process the inflow at different rates due to the established equipment's dissimilarity (dissimilarity rooted in efficiency, function type, etc.). On the other hand, these variable rates are the principal reason for forming a queue at each server. So, the limitation of service rates engenders congestion at hub nodes leading to a queue. On the other hand, a maximum queue length is assumed at each hub which avoids server collapse. Therefore, the proposed model in this paper aims to provide the most optimal set of hub nodes to minimize the costs (including the establishment of hubs and transportation costs) while considering the implied constraints on each hub's queue system. It may seem conspicuous to determine extra nodes as hubs or even equip the servers of each hub nodes to deter forming a queue at these nodes. The discount rates of flow between hubs may even bolster the idea of more hubs establishment. However, when it comes to the foundation huge expenses, it is the total costs of the model that

2 | Literature Review

Three coverage criteria were proposed by Campbell [1]. The origin-destination pair (i, j) is covered by hubs *k* and *m* if:

1. The cost from *i* to *j* via *k* and *m* does not surpass a given quantity.
2. The cost of each arc from *i* to *j* via *k* and *m* does not surpass a given quantity.
3. Each of the origin–hub, and hub–destination arcs satisfies different given values.

The multiple-allocation and the uncapacitated single-allocation problems were studied decades ago, and it was Campbell [1] who proposed the first mathematical model for the multiple-allocation problem. Later, O'Kelly [2] developed several models for hub location problems. They modeled the organization of a single and two hub networks. Klincewicz [3] proposed an effective algorithm for the uncapacitated hub location problem. Skorin-Kapov et al. [4], Ernst and Krishnamoorthy [5], and Mayer and Wagner [6] proposed multiple advances to the hub location problem, while Hamacher et al. [7] and Marı'n et al. [8] studied the

In Press

problem in a polyhedral manner. More advances have been presented by Marı'n (2005a) [9] and Canovas et al. [10]. O'Kelly [11], Klincewicz [12], Skorin-Kapov et al. [4], Aykin [13], and Ernst and Krishnamoorthy [5] studied single-allocation. Furthermore, Aykin [14], Ebery et al. [15], Campbell, [1], Boland et al. [16], and Marı'n (2005b) [17] studied the capacitated multiple-allocation problem. The capacitated single-allocation problem has also been studied by Ernst and Krishnamoorthy [18], Labbé et al. [19], Contreras et al. [20], [21], and Contreras [22]. Studies of Campbell et al. [23], Alumur and Kara [24], Ebery [25], Adler and Hashai [26], and O'Kelly and Bryan [27] will be extensive sources for eager readers.

Riedi et al. [28] proposed a model for multi-scale queuing (MSQ). Ashour and Le-Ngoc [29] developed an MSQ model for variable-service rate multi-scale queuing (VS-MSQ) to evaluate priority queues. They presented an analytical framework to estimate the length of the queue and delay survivor functions for a priority queuing system with varying service rates. Marianov and Serra [30] considered congestion in the network and proposed a mathematical model to find optimal locations of hub nodes. They considered the most congested airports and modeled them as M/D/c queuing systems. To solve the proposed model, they linearized the probabilistic constraint and employed a tabu search algorithm. Elhedhli and Xiaolong [31] modeled the congestion effect at a particular hub utilizing a convex cost function that increases exponentially as more flows are directed through that hub. Mohammadi et al. [32] modeled hubs as the most crowded network parts, as M/M/c queuing systems. Rahimi et al. [33] considered congestion and uncertainty in the hub simultaneously to design a network with multiple objective functions. In their study, Zhalechian et al. [34] focused on social responsibility and congestion in the hub and spoke problem. Khodemoni-yazdi et al. [35] studied hierarchical hub location problems applying a two-objective model while incorporating the queuing system in their study. Karimi-Mamaghan et al. [36] proposed a novel bi-objective model to consider congestion in both hubs and hub-to-hub connections in a hub-and-spoke network. The goal of the model was to minimize the time and cost of transportation. They also developed a new model based on the GI/G/c queuing system to study the congestion in the bub nodes, while a traffic model was utilized for the flow congestion between hubs.

Bütün et al. [37] developed a model to optimize the hub-and-spoke network for the shipping sector under congestion in the hub nodes. In their capacitated directed cycle hub location and cargo routing problem, the nonlinear costs of congestion were then linearized using a linear approximation. Alumur et al. [38] mentioned that congestion in hub the HLP will be a significant issue for passenger and cargo delivery due to the extra costs it imposes on the model that usually makes the hub location models nonlinear. Najy and Diabat [39] studied a multiple-allocation uncapacitated HLP that incorporates economies of scale and congestion in nodes. They utilized a Benders decomposition approach to solve their proposed model. Alumur et al. [40] modeled the congestion in the hubs and proposed a measure for congestion value in the HLP. They showed that ignoring congestion in the hubs can even make the hub model infeasible.

In Press

Several studies in the HLP have been conducted to cover some novel aspects of the HLP in recent years, which are mentioned here. Sadeghi et al. [41] regarded the budget for travel time in p-hub covering problems. Environmental aspects of the HLP were studied by Zhalechian et al. [42] by proposing a mathematical model involving noise pollution of the transportation facilities. Korani and Eydi [43] proposed a bi-level programming model whose first level was to minimize the cost of establishing a hub network, and in the second level, the service level loss was reduced. They solved the model by developing a two-stage penalty heuristic method utilizing penalty functions. Mahmoodjanloo et al. [44] utilized a bi-level programming model to consider the customer loyalty of a transportation company and minimize the transportation pricing by designing a competitive hub. They tackled the hubs' location and assignment decisions first, while pricing decisions were considered next. To solve the model, they utilized a scatter search algorithm and a metaheuristic method. Golestani et al. [45] focused on tackling an HLP for transportation of multiple perishable goods in a cold supply chain whose maintenance temperatures were not the same. They proposed a bi-objective model to minimize the hub establishment, transportation, and CO2 costs while adjusting the storage temperatures. Further studies in this regard can be reached in Alizadeh Firozi et al. [46], which has focused on solving a single-allocation HLP by an improved genetic algorithm; Zahiri et al. [47], Farrokhi-Asl and Tavakkoli-Moghaddam [48], and Ghodratnama et al. [49].

In the hub location problem literature, multi-server hubs with different service rates are not addressed, and this was our primary motivation to incorporate this real-world characteristic of the problem into the model. In this paper, we obtain the distribution rate of each node. The service capacity of each node should be studied to determine the probability distribution of the service rates. Having the service rates determined, other variables, including the queue length, waiting time, and system time, can be achieved using mathematical queueing equations. Since each node has the role of collector and distributor, each hub has a service system that includes one server or more. Given each hub's arrival and service rates, the whole system can be analyzed as a queuing system. In the proposed system, each hub node comprises several servers that handle the inflow to the hub node. Each server has a service rate, which depends on the system status and the number of customers in the system.

An example is given here to illustrate the proposed model in this paper. Suppose that several cities should be interconnected owing to their mutual demand and supply interactions. So, the cities are considered nodes (Figure 1), and if a city in this network is selected as a hub, it can have multiple servers to process the flow from/to other cities. Therefore, the hub median model with multiple allocation modes can be used in this example. By selecting a specific city as a hub, some facilities should be established, imposing an establishment cost on the model, which is added to the transportation costs of the model. If l is the number of servers in a hub city, then the city is assumed to be assigned with β_l % of the whole load. The service rate follows an exponential pattern in the cargo service centers at each hub while remaining finite. On the other hand, while the number of customers in each hub is desired to be less than a given value, this will not always be the case. Furthermore, the output of each server is identical in every respect, and each server offers the same service quality as other servers. Consequently, in this instance, in the presence of service rate limitation and the excess number of customers in the system, it will be inevitable for a queue to be formed at each hub.

In Press

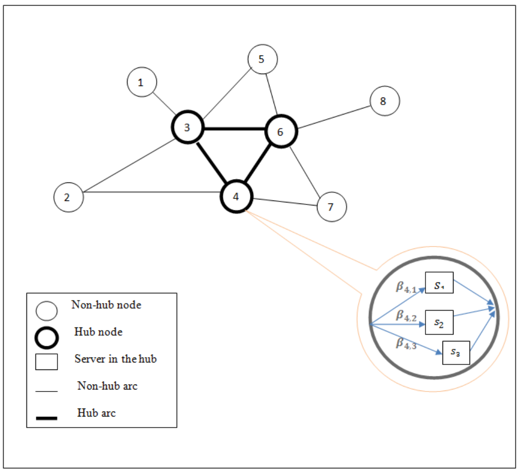

Fig. 1. The schematic form of the proposed approach.

3 | Modelling

This paper addresses a hub location problem based on the p-hub median model that considers the limited number of customers in the queue for each hub, and the objective is to minimize the total costs, including transportation and establishments costs. P-hub median models use three-part paths for collection, transfer, and distribution. Campbell [1] considers the cost for an origin-destination path as C_{ijkm}=d_{ik}+αd_{km}+d_{mj} where d_{km} is the distance between two hubs (*k* and *m*), and d_{ik} and d_{mj} depict the distances between the hub and non-hub nodes. The parameter α ($0 \le \alpha \le 1$) represents the discount amount corresponding to hub arcs to reflect the lower transportation cost due to the higher transportation scales.

The variables and parameters of the mathematical model are as follows:

Decision variables:

X_{ijkm}: The fraction of flow from node *i* (origin) to node *j* (destination) from path *i-k-m-j*

$$Y_k : \begin{cases} 1 & \text{if node } i \text{ is allocated to hub } k; \\ 0 & \text{otherwise,} \end{cases}$$

Parameters:

C_{ijkm}: The transportation cost from node *i* to node *j* from path *i-k-m-j*

W_{ij} : The flow (e.g., freight volume) to be transported from node *i* to node *j*.

F_k : The fixed cost of opening a hub at node *k*.

F_{km}: The fixed cost related to the establishment of an arc between hub nodes *k* and *m*.

P: The number of hubs.

$\theta_{q,kl}$: The maximum acceptable value for the probability of an excessive queue length at the l^{th} server of hub *k*.

b_{kl}: Maximum acceptable queue length at the l^{th} server of hub *k*.

$$Min \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\sum_{k=1}^{n}\sum_{m=1}^{n} C_{ijkm} W_{ij} X_{ijkm} + \sum_{k=1}^{n} F_k Y_k + \sum_{k=1}^{n-1}\sum_{m=k+1}^{n} \frac{1}{2} F_{km} X_{kmkm} \quad (1)$$

Subject to:

$$\sum_{k=1}^{n}\sum_{m=1}^{n} X_{ijkm} = 1 \qquad \forall (i,j,k,m) \mid i < j \quad (2)$$

$$\sum_{k=1}^{n} Y_k = P \quad (3)$$

$$X_{ijkk} + \sum_{m \ne k} (X_{ijkm} + X_{ijmk}) \le Y_k \qquad \forall (i,j,k,m) \mid i < j \quad (4)$$

$$P\,[\text{queue length at the } l^{th} \text{ server of hub } k > b_{kl}] \le \theta_{q,kl} \quad \forall (k,l) \quad (5)$$

$$X_{ijkm} \ge 0 \qquad \forall (i,j,k,m) \quad (6)$$

$$Y_k \in \{0,1\} \qquad \forall (k) \quad (7)$$

In Press

$$0 \leq \theta_{q,kl} \leq 1 \qquad \forall(k,l) \qquad (8)$$

The objective function (1) minimizes the total cost, including the variable costs (collection, distribution, and transportation) and fixed costs (hub and arc establishment). Constraint (2) ensures that each origin-destination flow goes through at least one hub node. The selection of exactly p hubs is guaranteed in constraint (3). Constraint (4) ensures that the flow between every two nodes should pass through at least one hub. Constraint (5) demonstrates that the probability of exceeding the maximum acceptable queue length in the l^{th} server of hub *k* is, at most, equal to a pre-defined maximum value at this hub. Constraints (6) define variables' sign and acceptable value, and constraint (8) shows the upper and lower bound of θ.

It is assumed that the service rate is in each server is a function of the system status. Therefore, by changing the system status, the service rate is changed. In this paper, it is assumed that the relationship between service rate and the system status is as follows:

$$\mu_n = n^c . \mu \qquad (9)$$

Where

μ is the average system rate when one client is in the system.

μ_n is the average system rate when *n* clients are in the system.

n^c is a constant factor that represents the state dependency of the service rate.

Also, it is assumed that the input rate follows a Poisson process with parameter λ . Thus, the arrival rate is as follows:

$$\lambda_n = \lambda \qquad (10)$$

As mentioned before, each node consists of some servers, and the servers function independently. The customers are allowed to choose each server on each node and wait in its queue. The arrival rate of each node *k* is λ_k which follows a Poisson process, and the arrival rate of server *l* at node *k*, i.e. λ_{kl} is calculated as follows:

$$\lambda_{kl} = \beta_{kl} . \lambda_k \qquad \forall(k,l) \qquad (11)$$

And

$$\sum_{l=1}^{L} \beta_{kl} = 1 \qquad \forall(k) \qquad (12)$$

The probabilistic equation (5) needs to be rewritten in an analytic form to make the mathematical model solvable. In order to obtain a deterministic linear constraint corresponding to this equation, p_s is defined as the steady-state probability of existing s customers in the system with one server. Then, *Eq.* (5) can be written as:

In Press

$$\sum_{s=b_{kl}+2}^{\infty} p_s \leq \theta_{q,kl} \quad or \quad 1-\sum_{s=0}^{b_{kl}+1} p_s \leq \theta_{q,kl} \tag{13}$$

While both forms of *Eq.* (13) convey the same concept, the left-hand side in either form illustrates the probability of exceeding the maximum acceptable clients waiting in a queue at the l^{th} server of hub k.

In the forthcoming steps, p_s will be supplanted by an equivalent expression to be substituted in the primary model.

The arrival and service rates of the system are assumed to be λ and μ respectively so, the arrival and service rates of each state will be $\lambda_n = \lambda$ and $\mu_n = n^c.\mu$, respectively.

Also, for the considered queueing system in this paper (i.e., an exponential model with variable service rates), the probability of the presence of no customer in the hubs is:

$$p_0 = \left[1 + \sum_{n=1}^{\infty} \frac{\left(\frac{\lambda}{\mu}\right)^n}{(n!)^c}\right]^{-1} \tag{14}$$

The probability of existing n customers in the system is:

$$p_n = \left(\frac{\lambda}{\mu}\right)^n \frac{1}{(n!)^c} p_0 \tag{15}$$

As there are infinite numbers of terms in *Eq.* (13), we propose a method to simplify the computation of p_0. In this method, we define *M* as a big number and replace the ∞ with *M*. Thus, the p_0 can be written as:

$$p_0 = \left[1 + \sum_{n=1}^{M} \frac{\left(\frac{\lambda}{\mu}\right)^n}{(n!)^c}\right]^{-1} \tag{16}$$

The more quantity of *M*, the more accuracy in computing p_0. We define ε as a minimum required accuracy for p_0.

$$\left[1 + \sum_{n=1}^{s-1} \frac{\left(\frac{\lambda}{\mu}\right)^n}{(n!)^c}\right]^{-1} - \left[1 + \sum_{n=1}^{s} \frac{\left(\frac{\lambda}{\mu}\right)^n}{(n!)^c}\right]^{-1} < \varepsilon \tag{17}$$

Then we consider $M = s$ to compute p_0.

$$p_0 = \left[1 + \sum_{n=1}^{s} \frac{\left(\frac{\lambda}{\mu}\right)^n}{(n!)^c}\right]^{-1} \tag{18}$$

Thus, the constraint (13) and consequently the constraint (5) can be rewritten as *Eq.* (19):

In Press

$$\sum_{n=0}^{b_{kl}+1}\left(\frac{\lambda}{\mu}\right)^{n}\frac{1}{(n!)^{c}}\left[1+\sum_{n=1}^{s}\frac{\left(\frac{\lambda}{\mu}\right)^{n}}{(n!)^{c}}\right]^{-1}\geq 1-\theta_{q,kl} \qquad (19)$$

Turning to the arrival rate to the hub *k*, *Eq.* (20) shows this value for the hub *k* as follows:

$$\lambda_k=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n}a_{ij}\,x_{ijkm}+\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n}a_{ij}\,x_{ijmk}\right) \qquad \forall(k) \qquad (20)$$

Furthermore, the arrival rate to the l^{th} server of hub *k* is given by:

$$\lambda_{kl}=\beta_{kl}\times\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n}a_{ij}\,x_{ijkm}+\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n}a_{ij}\,x_{ijmk}\right) \qquad \forall(k,l) \qquad (21)$$

Where a_{ij} is the peak hour transportation flow from node *i* to node *j* by using hub *k*. In the meantime, the total arrival rate to the hub *k* is the sum of arrival rates from origin *i* and hub *m*.

According to Marianov and Serra [30], we can numerically solve *Eq.* (19) and find the maximum value of variable λ indicated by λ_{max}. All the possible values for the λ that are less than λ_{max} can satisfy *Eq.* (19). Therefore, *Eq.* (19) is equivalent to the following equation:

$$\lambda\leq\lambda_{max} \qquad (22)$$

If the value of λ_{max} is computed for each server of node *k* (assuming that there is a difference between nodes in terms of service time), we could rewrite $\lambda_{kl}\leq\lambda_{max,kl}$ as follows:

$$\beta_{kl}\times\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n}a_{ij}\,x_{ijkm}+\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{m=1}^{n}a_{ij}\,x_{ijmk}\right)\leq\lambda_{max,kl} \qquad (23)$$

To solve the model, the following steps (that are proposed by Marianov and Serra [30]) are as follows:

1. Estimating the service rate for each server of hubs.
2. Finding the $\lambda_{max,kl}$ for each server of hubs.
3. Solving the following model by using the value of $\lambda_{max,kl}$:

In Press

$$Min \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{k=1}^{n} \sum_{m=1}^{n} C_{ijkm} W_{ij} X_{ijkm} + \sum_{k=1}^{n} F_k Y_k + \sum_{k=1}^{n-1} \sum_{m=k+1}^{n} \frac{1}{2} F_{km} X_{kmkm} \quad (1)$$

Subject to:

$$\sum_{k=1}^{n} \sum_{m=1}^{n} X_{ijkm} = 1 \qquad \forall (i,j,k,m) \mid i < j \quad (2)$$

$$\sum_{k=1}^{n} Y_k = P \quad (3)$$

$$X_{ijkk} + \sum_{m \neq k} (X_{ijkm} + X_{ijmk}) \leq Y_k \qquad \forall (i,j,k,m) \mid i < j \quad (4)$$

$$\beta_{kl} \times \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{m=1}^{n} a_{ij}\, x_{ijkm} + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{m=1}^{n} a_{ij}\, x_{ijmk} \right) \leq \lambda_{max,kl} \quad (23)$$

$$X_{ijkm} \geq 0 \qquad \forall (i,j,k,m) \quad (6)$$

$$Y_k \in \{0,1\} \qquad \forall (k) \quad (7)$$

$$0 \leq \theta_{q,kl} \leq 1 \qquad \forall (k,l) \quad (8)$$

4 | Numerical Results

To evaluate our proposed approach, we conduct a numerical experiment on the CAB dataset that includes the distance and flow (i.e., intercity passengers) among 25 cities in the United States [27] portrayed in *Fig. 2*. Several instances are designed based on different values of transportation cost discount rates (α =0.2, 0.5, and 0.8), the number of hubs (p=4, 8, and 12), and the acceptable number of customers in the queue (b= 5, 20).

Fig. 2. CAB data set illustrating 25 cities in the United States.

In the generated sample problems, a fixed cost was considered for establishing hub nodes and hub-to-hub arcs. It is assumed that each hub includes three servers with different service rates. The service rates of each server is a function of the number of people in the queue ($c = 0.2$, $\mu_{kln} = \mu_{kl} * n^{0.2}$). Another assumption was that the customers entrance to the first, second, and third server with β_{k1} =0.5, β_{k2}=0.2 and β_{k3}=0.3 respectively. For each server, the value of λ_{kl} was calculated based on the $\theta_{q,kl} = 0.95$, and was imported to the model. The model was coded and solved in GAMS 22.2 optimization software. The results of the solutions (the values of objective function and the nodes selected as hubs) corresponding to each problem are represented in *Table 1*.

In press

Table 1. The resulting solutions for designed problems.

α	p	b_{kl}	Objective Function's Value	Hub nodes
0.2	4	5	5332	8, 14, 15, 17
		20	4335	1, 8, 17, 20
	8	5	3361	2, 4, 6, 12, 13, 14, 17, 23
		20	3061	2, 4, 6, 12, 13, 14, 17, 23
	12	5	3412	1, 2, 4, 6, 8, 13, 14, 15, 17, 21, 22, 23
		20	3344	1, 2, 4, 6, 7, 13, 14, 16, 17, 19, 21, 23
0.5	4	5	5730	8, 14, 15, 17
		20	4825	4, 8, 17, 20
	8	5	4149	1, 4, 12, 13, 14, 17, 20, 23
		20	4037	1, 4, 12, 13, 14, 17, 20, 23
	12	5	4269	1, 2, 4, 6, 13, 14, 16, 17, 19, 21, 22, 23
		20	4252	1, 2, 4, 6, 13, 14, 16, 17, 19, 21, 22, 23
0.8	4	5	5831	8, 14, 15, 17
		20	5105	4, 8, 17, 20
	8	5	4833	1, 4, 14, 16, 17, 19, 20, 23
		20	4780	1, 4, 14, 17, 19, 20, 21, 23
	12	5	5084	1, 2, 4, 6, 13, 14, 15, 16, 17, 19, 21, 23
		20	4952	1, 2, 4, 6, 13, 14, 15, 16, 17, 19, 21, 23

According to *Table 1*, by increasing the value of α, the total cost increases. Also, it is observed that the total cost for $p = 8$ is less than the total cost for $p = 4$ and $p = 12$. This means that if the number of hubs is increased from 4 to 8, the transportation cost will decrease because the distances between non-hub and hub nodes will be reduced, and the number of arcs with a discounted fee will increases. In such a situation, the transportation cost saving is more than the establishment costs for extra hubs. Conversely, by increasing the number of hubs from 8 to 12, the establishment cost would be higher than the transportation cost saving. Therefore, a network with eight hubs is more cost-efficient than a network with 4 or 12 hubs. The critical point is that if the maximum acceptable value for the queue length is altered from 5 to 20, the total cost decreases because the problem is less restricted. In more detail, in the cases with lower allowable queue length, the model is forced to reduce the number of customers in the queue by assigning higher service rates to the hubs (which implies higher establishment fixed costs to the model). In most solutions, increasing the allowable length of the queue does not change the selected hubs; it changes the assignment of non-hub to hub nodes, however, leading to a reduced total cost (for instance, the problem with $p = 8$, $\alpha = 0.5$ and b_{kl} altering from 5 to 20).

The results of two computational solutions are shown in *Fig.* 3 and *Fig. 4*. *Fig. 3* indicates the hub nodes and hub arcs for the problem with $p = 4$, $\alpha = 0.2$, and $b_{kl} = 5$. *Fig. 4* indicates the hub nodes and hub arcs for the following parameters: $p = 4$, $\alpha = 0.2$, and $b_{kl} = 20$.

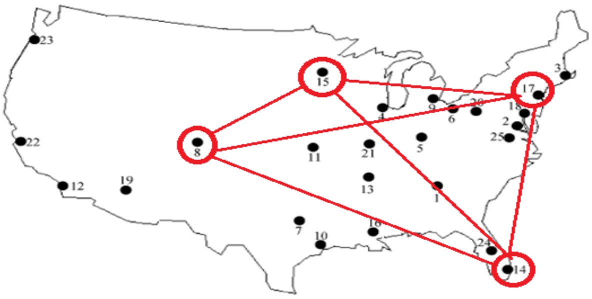

Fig. 3. Hub nodes and related arcs for a problem with $p = 4$, $\alpha = 0.2$ and $b_{kl} = 5$.

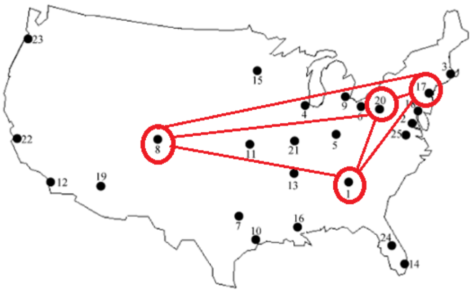

Fig. 4. Hub nodes and related arcs for a problem with $p = 4$, $\alpha = 0.2$ and $b_{kl} = 20$.

4 | Conclusion and Future Research

In this paper, a p-hub median problem was addressed, in which several servers at each hub provide service to the customers with limited service rates so, a queue with a confined queue length will be created at each hub. The goal is to restrict the queue length to prevent congestion in the hubs. We proposed a mixed-integer linear programming (MILP) model to minimize the total costs, including transportation and fixed establishment costs, by considering the queue length constraint at each hub. A numerical experiment based on the CAB dataset was conducted to evaluate the proposed approach. The numerical experiment results indicated the impact of restricting the queue length on the solutions by changing the total costs and, in some cases, changing the selected hub nodes. For example, decreasing the allowable queue length increases the total cost due to establishing more equipped hubs with higher service rates and establishment costs. Consequently, strategic decisions regarding how many nodes should act as hubs and how many customers are allowed to receive service at each hub server impact the total costs of the model.

To provide a guideline for future research, the proposed model in this paper can be extended to involve more real-world conditions such as environmental, social, and energy considerations when establishing a hub. Then, if such constraints were added to the model, the model's complexity would undoubtedly increase, making the researchers inevitable of utilizing metaheuristics methods to tackle such complexity.

Conflicts of Interest

All co-authors have seen and agree with the manuscript's contents, and there is no financial interest to report. We certify that the submission is original work and is not under review at any other publication.

References

[1] Campbell, J.F. (1994). Integer programming formulations of discrete hub location problems. *European Journal of Operational Research, 72,* 387–405. https://doi.org/10.1016/0377-2217(94)90318-2

In Press

[2] O'Kelly M.E. (1986). The location of interacting hub facilities. *Transportation Science, 20,* 92–106. https://doi.org/10.1287/trsc.20.2.92

[3] Klincewicz, J.G. (1996). A dual algorithm for the uncapacitated hub location problem. *Location Science*, *4*, 173–184. https://doi.org/10.1016/S0966-8349(96)00010-1

[4] 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. https://doi.org/10.1016/0377-2217(95)00100-X

[5] Ernst, A.T., & Krishnamoorthy, M. (1998). An exact solution approach based on shortest-paths for p-hub median problems. *INFORMS Journal on Computing*, *10*(2), 149–162. https://doi.org/10.1287/ijoc.10.2.149

[6] Mayer, G., & Wagner, B. (2002). Hub Locator: An exact solution method for the multiple-allocation hub location problems. *Computers and Operations Research*, *29*(6), 715–739. https://doi.org/10.1016/S0305-0548(01)00080-6

[7] Hamacher, H.W., Labbé, M., Nickel, S., & Sonneborn, T. (2004). Adapting polyhedral properties from facility to hub location problems. *Discrete Applied Mathematics*, *145*(1), 104–116. https://doi.org/10.1016/j.dam.2003.09.011

[8] Marín, 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. https://doi.org/10.1016/j.ejor.2004.09.047

[9] Marín, A. (2005a). Uncapacitated Euclidean hub location: Strengthened formulation new facets and a relax-and-cut algorithm. *Journal of Global Optimization*, *33*, 393–422. https://doi.org/10.1007/s10898-004-6099-4

[10] Cánovas, L., García, S., & Marín, A. (2007). Solving the uncapacitated multiple-allocation hub location problem by means of a dual ascent technique. *European Journal of Operational Research, 179*(3), 990–1007. https://doi.org/10.1016/j.ejor.2005.08.028

[11] 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. https://doi.org/10.1016/S0377-2217(87)80007-3

[12] Klincewicz, J.G. (1991). Heuristics for the p-hub location problem. *European Journal of Operational Research, 53*(1), 25–37. https://doi.org/10.1016/0377-2217(91)90090-I

[13] Aykin, T. (1995). Networking policies for hub-and-spoke systems with application to the air transportation system. *Transportation Science 29*(3), 201–221. https://doi.org/10.1287/trsc.29.3.201

[14] Aykin, T. (1994). Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem. *European Journal of Operational Research*, *79*(3), 501–523. https://doi.org/10.1016/0377-2217(94)90062-0

[15] Ebery, J., Krishnamoorthy, M., Ernst, A., & Boland, N. (2000). The capacitated multiple-allocation hub location problems: Formulations and algorithms. *European Journal of Operational Research, 120*(3), 614–631. https://doi.org/10.1016/S0377-2217(98)00395-6

[16] 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. https://doi.org/10.1016/S0377-2217(03)00072-9

[17] Marín, A. (2005b). Formulating and solving the splittable capacitated multiple-allocation hub location problems. *Computers and Operations Research*, *32*(12), 3093–3109. https://doi.org/10.1016/j.cor.2004.04.008

[18] Ernst, A.T., & Krishnamoorthy, M. (1999). Solution algorithms for the capacitated single allocation hub location problem. *Annals of Operations Research*, *86*, 141–159. https://doi.org/10.1023/A:1018994432663

[19] Labbé, M., Yaman, H., & Gourdin, E. (2005). A branch and cut algorithm for hub location problems with single assignment. *Mathematical Programming*, *102*, 371–405. https://doi.org/10.1007/s10107-004-0531-x

[20] Contreras, I., Díaz, J.A., & Fernández, E. (2008). Branch and price for large-scale capacitated hub location problems with single assignment. *INFORMS Journal on Computing*, *23*(1), 41-55. https://doi.org/10.1287/ijoc.1100.0391

[21] Contreras, I., Díaz, J.A., & Fernández, E. (2009). Lagrangian relaxation for the capacitated hub location problem with single assignment. *Operations Research Spectrum*, *31*, 483–505. https://doi.org/10.1007/s00291-008-0159-y

[22] Contreras, I. (2009). *Network Hub Location: Models, algorithms and related problems*. Ph.D. thesis, Department of Statistics and Operations Research, Technical University of Catalonia, Spain. https://books.google.com/books?id=S4PKMwEACAAJ

[23] Campbell, J.F., Ernst, A.T., & Krishnamoorthy, M. (2002). Hub location problems. In: Drezner, Z., & Hamacher, H.W. (Eds.), *Facility Location: Applications and Theory* (pp. 373–407). Berlin Heidelberg: Springer.

In Press

[24] Alumur, S., & Kara, B.Y. (2008). Network hub location problems: The state of the art. *European Journal of Operational Research, 190* (1), 1–21. https://doi.org/10.1016/j.ejor.2007.06.008

[25] Ebery, J. (1999). Solving large single-allocation p-hub problems with two or three hubs. *European Journal of Operational Research, 128*(2), 447-458. https://dx.doi.org/10.1016/S0377-2217(99)00370-7

[26] Adler, N., &Hashai.N. (2005). Effect of open skies in the Middle East region. *Transportation Research Part A: Policy and Practice, 39*(10), 878–894. https://doi.org/10.1016/j.tra.2005.04.001

[27] O'Kelly, M.E., Bryan, D., Skorin-Kapov, D., & Skorin-Kapov, J. (1996). Hub network with single and multiple allocation: a computational study. *Location Science, 4*(3), 125–38. https://doi.org/10.1016/S0966-8349(96)00015-0

[28] Riedi, R.H., Crouse, M.S., Ribeiro, V.J., & Baraniuk, R.G. (1999). A multiracial wavelet model with application to network traffic. *IEEE Transactions on Information Theory, 45*(3), 992-1018. https://doi.org/10.1109/18.761337

[29] Ashour, M., & Le-Ngoc, T. (2003). Priority Queuing of Long Range Dependent Traffic. *GLOBECOM '03. IEEE Global Telecommunications Conference (IEEE Cat. No.03CH37489*), 6, 3025–3029. https://doi.org/10.1109/GLOCOM.2003.1258791

[30] Marianov, V., & Serra, D. (2003). Location models for airline hubs behaving as M/D/c queues. *Computers and Operations Research 30*(7), 983–1003. https://doi.org/10.1016/S0305-0548(02)00052-7

[31] Elhedhli, S., & Xiaolong Hu, F. (2005). Hub-and-spoke network design with congestion. *Computers & Operations Research, 32*(6), 1615–1632. https://doi.org/10.1016/j.cor.2003.11.016

[32] Mohammadi, M., Jolai, F., & Rostami, H. (2011). An M/M/c queue model for hub covering location problem. *Mathematical and Computer Modelling, 54*(11-12), 2623–2638. https://doi.org/10.1016/j.mcm.2011.06.038

[33] Rahimi, Y., Tavakkoli-Moghaddam, R., Mohammadi, M., & Sadeghi, M. (2016). Multi-objective hub network design under uncertainty considering congestion: An *M/M/c/K* queue system. *Applied Mathematical Modelling, 40*(5-6), 4179–4198. https://doi.org/10.1016/J.APM.2015.11.019

[34] Zhalechian, M., Tavakkoli-Moghaddam, R., & Rahimi, Y. (2017). A self-adaptive evolutionary algorithm for a fuzzy multi-objective hub location problem: An integration of responsiveness and social responsibility. *Engineering Applications of Artificial Intelligence, 62*, 1-16. https://doi.org/10.1016/j.engappai.2017.03.006

[35] Khodemoni-yazdi, M., Tavakkoli-Moghaddam, R., Bashiri, M., & Rahimi, Y. (2019). Solving a new bi-objective hierarchical hub location problem with an M/M/c queuing framework. *Engineering Applications of Artificial Intelligence, 78*, 53-70. https://doi.org/10.1016/j.engappai.2018.10.004

[36] Karimi-Mamaghan, M., Mohammadi, M., Pirayesh, A., Karimi-Mamaghan, A.M., & Irani, H. (2020). Hub-and-spoke network design under congestion: A learning based metaheuristic. *Transportation Research Part E: Logistics and Transportation Review, 142*. https://doi.org/10.1016/j.tre.2020.102069

[37] Bütün, C., Petrovic, S., & Muyldermans, L. (2021). The capacitated directed cycle hub location and routing problem under congestion. *European Journal of Operational Research, 292*(2), 714-734. https://doi.org/10.1016/j.ejor.2020.11.021

[38] Alumur, S.A., Campbell, J.F., Contreras, I., Kara, B.Y., Marianov, V., & O'Kelly, M.E. (2021). Perspectives on modeling hub location problems. *European Journal of Operational Research, 291*(1), 1-17. https://doi.org/10.1016/j.ejor.2020.09.039

[39] Najy, W., & Diabat, A. (2020). Benders decomposition for multiple-allocation hub-and-spoke network design with economies of scale and node congestion. *Transportation Research Part B: Methodological, 133*, 62-84. https://doi.org/10.1016/j.trb.2019.12.003

[40] Alumur, S.A., Nickel, S., Rohrbeck, B., & Saldanha-da-Gama, F. (2018). Modeling congestion and service time in hub location problems. *Applied Mathematical Modelling, 55*, 13-32. https://doi.org/10.1016/j.apm.2017.10.033

[41] Sadeghi, M., Tavakkoli-Moghaddam, R., & Babazadeh, R. (2018). An efficient artificial bee colony algorithm for a reliable p-hub covering location problem with travel time budget. *International Journal of Industrial Engineering: Theory, Applications and Practice, 25* (1), 40–53.

[42] Zhalechian, M., Tavakkoli-Moghaddam, R., Rahimi, Y., & Jolai, F. (2017). An interactive possibilistic programming approach for a multi-objective hub location problem: Economic and environmental design. *Applied Soft Computing, 52*, 699-713. https://doi.org/10.1016/j.asoc.2016.10.002

In Press

[43] Korani, E., & Eydi, A. (2021). Bi-level programming model and KKT penalty function solution approach for reliable hub location problem. *Expert Systems with Applications*, *184*. https://doi.org/10.1016/j.eswa.2021.115505

[44] Mahmoodjanloo, M., Tavakkoli-Moghaddam, R., Baboli, A., & Jamiri, A. (2020). A multi-modal competitive hub location pricing problem with customer loyalty and elastic demand. *Computers & Operations Research*, *123*. https://doi.org/10.1016/j.cor.2020.105048

[45] Golestani, M., Moosavirad, S.H., Asadi, Y., & Biglari, S. (2021). A Multi-Objective Green Hub Location Problem with Multi Item-Multi Temperature Joint Distribution for Perishable Products in Cold Supply Chain. *Sustainable Production and Consumption*, *27*, 1183-1194. https://doi.org/10.1016/j.spc.2021.02.026

[46] Alizadeh Firozi, M., Kiani, V., & Karimi, H. (2021). Improved Genetic Algorithm with Diversity and Local Search for Uncapacitated Single Allocation Hub Location Problem. *Journal of Decisions and Operations Research*. http://dx.doi.org/10.22105/dmor.2021.272989.1325

[47] Zahiri, B., Mousazadeh, M., & Bozorgi-Amiri, A. (2014). A Robust Stochastic Programming Approach for Blood Collection and Distribution Network Design. *International Journal of Research in Industrial Engineering*, *3*(2), 1-11.

[48] Farrokhi-Asl, H., & Tavakkoli-Moghaddam, R. (2016). Solving a Bi-Objective Transportation Vehicle Routing Problem for Waste Collection by a New Hybrid Algorithm. *International Journal of Research in Industrial Engineering*, *5*(1(4)), 16-42. https://dx.doi.org/10.22105/riej.2017.49168

[49] Ghodratnama, R., Tavakkoli-Moghaddam, R., & Azaron, A. (2015). Robust and fuzzy goal programming optimization approaches for a novel multi-objective hub location-allocation problem: A supply chain overview. *Applied Soft Computing*, *37*, 255-276. https://doi.org/10.1016/j.asoc.2015.07.038

In Press