Computers & Operations Research 65 (2016) 174–188

Contents lists available at ScienceDirect

Computers & Operations Research

journal homepage: www.elsevier.com/locate/caor

The impact of hub failure in hub-and-spoke networks: Mathematical formulations and solution techniques

Nader Azizi [a,*], Satyaveer Chauhan [b], Said Salhi [a], Navneet Vidyarthi [b]

[a] Centre for Logistics and Heuristic Optimisation, Kent Business School, University of Kent, Canterbury, Kent CT27PE, UK
[b] John Molson School of Business, Concordia University, Montreal, Quebec, Canada H3G 1M8

ARTICLE INFO

Available online 27 May 2014

Keywords:
Hub-and-spoke network
Evolutionary algorithm
Hub failure
Logistics
Backup facilities

ABSTRACT

Hub facilities are subject to unpredictable disruptions caused by severe weather condition, natural disasters, labor dispute, and vandalism to cite a few. Disruptions at hubs result in excessive transportation costs and economic losses as customers (demand) initially served by these facilities must now be served by other hubs. In this study, we first present a novel mathematical model that builds hub-and-spoke systems under the risk of hub disruption. In developing the model, we assume that once a hub stops normal operations, the entire demand initially served by this hub is handled by a backup facility. The objective function of the model minimizes the weighted sum of transportation cost in regular situation and the expected transportation cost following a hub failure. We adopted a linearization for the model and present an efficient evolutionary approach with specifically designed operators. We solved a number of small problem instances from the literature using CPLEX for our enhanced mathematical model. The obtained results are also used as a platform for assessing the performance of our proposed meta-heuristic which is then tested on large instances with promising results. We further study and provide results for the relaxed problem in which demand points affected by disruption are allowed to be reallocated to any of the operational hubs in the network.

1. Introduction

The classical hub location problem deals with locating hub facilities and allocating demand to hubs to direct the flow between origin–destination pairs. In the hub location literature, it is common to assume that there is a link between every hub pair, there is no direct path between non-hub nodes, and there is economies of scale for using the inter-hub connections [2]. Depending on how non-hub nodes are allocated to hub facilities, two types of network are constructed namely single and multiple allocations. In the former, all the incoming and outgoing traffic of every node is transferred through a single hub, while in the latter each node in the network can receive and send flow through more than one hub. In this research, we focus on the single allocation p-hub median problem with hub unavailability consideration which we term SApHM-HU.

The hub location problem has various applications in the areas of transportation e.g., air passenger and cargo [5,28,29,35], less-than-truckload freight [11], rail freight [19], urban public transportation and rapid transit [31]. Other applications areas include postal delivery [15,8], express package and cargo delivery [24,39,3], telecommunications [22,7] and supply chains [25]. Hub-and-spoke systems have been the subject of many studies in the past three decades. O'Kelly [32,33] presented the first mathematical model for the single allocation p-hub median problem. Campbell [6] developed a linear integer formulation for the problem. Examples of other formulations that have been proposed in the literature include Ernst and Krishnamoorthy [15], Skorin-Kapov et al. [36], and Ebery [13]. The objective of the p-hub median problem is to determine the location of a predetermined number of facilities (p) and the allocation of the non-hubs to these open hubs such that the total transportation cost is minimized.

Traditional approaches to hub location problem assume that hub facilities are always available. In practice, however, one or more of these facilities may become unavailable from time to time due to, for example, weather conditions and/or natural disasters. To manage hub failure, two strategies are usually adopted in air transportation which include reactive (e.g., canceling, delaying, rescheduling, etc. [18]) and proactive strategies (e.g., investment in reliability improvement of existing facilities). Nevertheless, a disruption at a hub may significantly affects service level and result in excessive transportation cost as customers (demand) initially served by these facilities must now be served by other hubs.

* Corresponding author. Tel.: +44 1634 888941.
E-mail address: n.azizi@kent.ac.uk (N. Azizi).

http://dx.doi.org/10.1016/j.cor.2014.05.012
0305-0548/

1.1. Illustrating the impact of hub failure via an example

To evaluate the impact of hub failure on operating cost, we simulate service disruptions in a problem instance with 10 nodes and 3 hubs taken from the U.S. Civil Aeronautics Board which is known as CAB dataset [32]. The dataset is based on the airline passenger interactions between 25 US cities in 1970 and has been frequently used by hub location researchers. Fig. 1 illustrates the optimal network configuration for this problem where inter-hub discount factor i.e., α is 0.2. The total transportation cost of the network presented in Fig. 1 in regular situation is 491.93 units [36].

We assume that once a hub becomes unavailable, the flow initially passing through this facility is rerouted via one of the operating hubs in the network. In Fig. 1, for instance, if hub 6 (Cleveland) is disrupted then the entire flow that uses this hub as the first or the second hub in the path from origin i to destination j is rerouted via either hub 4 (Chicago) or hub 7 (Dallas-Fw). This rerouting strategy is important as in some applications a group of spokes need to be communicated via a single hub to which they are allocated. For instance, in postal service hub facilities are major sorting centres equipped with sorting machinery, optical recognition units, etc. These facilities provide a service to nearby regional offices. In such a system instead of assigning one vehicle between spoke-hub pair, a small fleet will operate for each hub region and each vehicle will visit a subset of cities on their own tours [8,30].

In the simulation, we examine three cases where one of the existing hubs in the network is assumed to be disrupted at a given time. The total network cost corresponding to each case, that includes the following cost elements, is then calculated. The first element of the resulting network cost is the *rerouting* cost of the flow through a backup facility. The rerouted flow initially uses the disrupted hub as either its first or second hub in the path from origin i to destination j. The second element is the *demand loss* cost that measures the cost of not meeting the demand at a disrupted facility (i.e., cost of the flow that either initiates or ends up at the disrupted hub). The third and the final element is the cost of transporting the flow between nodes that are not affected by the hub disruption (*routing* cost). The above three types of cost (i.e., routing, rerouting and demand loss costs) when summed up together make up the *new network cost*.

The total cost of the three new networks corresponding to the above three cases is summarized in Table 1 where "Min" sums the demand loss cost and the smaller of the two routing-rerouting costs associated with each scenario. Each of the two routing-rerouting costs in Table 1 corresponds to the case in which one of the operating hubs in the network is utilized as the backup facility for the disrupted hub. The lower of the two costs associated with the case where the most (economically) attractive rerouting path (i.e., the best backup facility) is utilized to maintain network operations; "Max" represents the network cost when the least attractive backup is utilized to transfer the flow.

Comparison of "Min" network costs for all three scenarios in Table 1 indicates that scenario 1 in which hub 6 is assumed disrupted and hub 4 is utilized as its backup has the lowest cost. The highest cost belongs to the case where hub 4 is assumed disrupted and any of the two other operating hubs in the network (i.e., hub 6 or hub 7) is utilized as the backup facility.

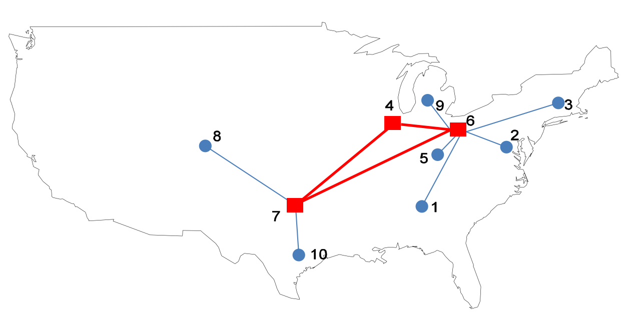
The results presented in Table 1 show that in the event of hub failure, deciding on backup facility largely affects the operating cost. For instance, in Case 1 if hub 7 is utilized as the backup for the assumed disrupted hub 6, the resulting network cost is estimated to be 1149.25 units. However, the network cost significantly reduces (709.41 unit) if hub 4 is used as the backup for hub 6 in the event of hub failure. Our results for the relatively small problem described above further indicates that hub failure causes an excessive cost which on average could increase the regular transportation cost by nearly 89%.

With regard to routing and rerouting costs, our analysis of the results presented in Table 1 shows that the most expensive hub in the network is hub 6. If disrupted, it will impose the largest amount of routing-rerouting cost to the system. This is understandable as more flow is transferred through this hub (i.e., hub 6) in comparison to the other two (hub) facilities in the network (see Fig. 1). One way to guard against such a scenario is to protect such a facility by increasing the level of security which obviously will require extra investment. There are however studies that incorporate these aspects into the modeling. Concerning the demand loss cost, results in Table 1 show that the most expensive facility is hub 4. This hub is the origin/destination of a significant amount of flow which is much higher than that in any of the two other hubs in the network. Therefore, the penalty for not meeting the demand in hub 4 is expected to be relatively high.

Using the data presented in Table 1, the lowest *Expected Transportation Cost* of the network is calculated by multiplying the minimum network costs (under column "Min" in Table 1) and

Table 1
New network costs following a single hub failure.

Disrupted hub	Backup hubs	Routing-rerouting cost	Demand loss cost	New network cost		
				Min	Max	Average
6	4	549.72	159.69	709.41	1149.25	929.33
	7	989.56				
4	6	317.94	540.48	858.43	858.43	858.43
	7	317.94				
7	4	484.80	273.01	757.81	808.92	783.37
	6	535.90				



1 Atlanta	2 Baltimore	3 Boston	**4 Chicago**	5 Cincinnati
6 Cleveland	**7 Dallas-Fw**	8 Denver	9 Detroit	10 Houston

Fig. 1. Optimal solution to a problem with 10 nodes and three hubs (CAB dataset) – single allocation. ■ Hub facility.

their corresponding hub failure probabilities. For instance, if probabilities of failures for hubs 6, 4, and 7 are 0.28, 0.15, and 0.1 respectively, then the expected transportation cost of the network would be 403.3 units.

In summary, the above analysis confirms that hub failure may largely increase the operating cost even if the most economically attractive backup facility is used to maintain network operations. It also shows that in the event of service disruption, the magnitude of excessive cost depends primarily on the amount of demand initially served by the disrupted hub and the selection of the backup facility.

1.2. A brief literature review

A number of papers addressed facility location problem with random disruptions. Few examples of such studies include Cui et al. [10], Li et al. [26], Peng et al. [34], Church et al. [9], Li et al. [27], Eiselt et al. [14], and Berman and Krass [4]. In this paper, our focus is on hub-and-spoke systems with possible hub failure for which a few models have been proposed in the literature [21]. Furthermore, a large portion of the literature on location problem concentrates on robustness to changes in demand or costs [37]. Our approach however, looks for robustness to changes in the supply network itself. This view of "network robustness" has some similarities with the work by Snyder and Daskin [37].

This study attempts to investigate a new methodology to design hub-and-spoke networks with "hub unavailability" consideration. More precisely, we propose models that build reliable networks by explicitly considering possible disruption at hub facilities at the design stage. Among others, our models aid to verify (1) if topologies of the networks with and without hub unavailability consideration are identical and (2) if failure cost could be reduced without a large increase in the regular transportation cost by incorporating hub unavailability as a part of the classical hub location model. While the first goal aims to determine whether the issues of hub unavailability and hub location-allocation could be dealt with separately, the second goal targets the amount of cost-saving that might be achieved by simultaneously addressing the two problems namely the hub location-allocation and the backup selection. To the best of our knowledge, our research is one of the first studies to consider hub unavailability in the context of hub-and-spoke systems.

The remainder of this paper is organized as follows. Problem description and model formulation including our linearization scheme are presented in Section 2. Section 3 describes the proposed evolutionary solution technique with its specific ingredients that we designed for the single allocation p-hub median problem with hub unavailability consideration. Computational results are reported in Section 4 followed by concluding remarks in Section 5.

2. Problem description and model formulation

In the single allocation p-hub median problem, each node is assigned to a particular hub and all the incoming and outgoing flows are routed through that hub. The total number of nodes in the network is assumed to be n, each node is considered a potential site and the number of hubs to open is p ($p<n$). A path from origin spoke i to destination spoke j includes three parts: collection from spoke i to the first hub k, transfer between first hub k and the second hub m, and distribution from hub m to destination j. The cost per unit flow along this route, $i \to k \to m \to j$, is calculated as $\chi \times c_{ik} + \alpha \times c_{km} + \delta \times c_{mj}$ where χ and δ are coefficients of collection and distribution respectively and α is the inter-hub discount factor. Let λ_{ij} be the amount of flow to be routed from origin i to destination j, the transportation cost from i to j routed via hubs k and m, C_{ikmj}, is then calculated as $C_{ikmj} = \lambda_{ij}(\chi \times c_{ik} + \alpha \times c_{km} + \delta \times c_{mj})$.

2.1. Notation

The three decision variables: hub location and allocation variable z, the route selection x, and the backup selection variable u are defined as follows:

$z_{ik} = 1$ if node i is assigned to hub k and $=0$ otherwise;
$x_{ikmj} = 1$ if flow from i to j passes through hub k and m and $=0$ otherwise;
$u_{kl} = 1$ if l is the backup for hub k and $=0$ otherwise.

Each hub k has a probability of failure q_k and we assume that only one hub will be disrupted at any given time. We further assume that hubs are uncapacitated.

In our formulation, FFC_k calculates the total cost of transporting flow from origin i to destination j via "the first hub" k, and BFC_m computes the total transportation cost from origin i to destination j through "the second hub" m. They are represented by the following expressions:

$$FFC_k = \sum_i \sum_m \sum_j C_{ikmj} x_{ikmj} \tag{1}$$

$$BFC_m = \sum_i \sum_{k \neq m} \sum_j C_{ikmj} x_{ikmj} \tag{2}$$

Given the above definition of C_{ikmj}, the regular transportation cost (RTC) is expressed as

$$RTC = \sum_i \sum_k \sum_m \sum_j C_{ikmj} x_{ikmj} \tag{3}$$

Using Eqs. (1) and (3), the RTC could be re-written as

$$RTC = \sum_k FFC_k \tag{4}$$

In a network with a single disrupted hub, three types of flows are identified:

(a) Flow that is initially planned to be transferred via the disrupted hub to reach their final destinations.
(b) Flow that is not affected by the disruption.
(c) Flow that originates from or end up at a disrupted hub.

For the type (a) flow, the disrupted hub is either its first or the second hub in the path from origin i to destination j. To maintain network operations, this type of flow (i.e., flow passing through a disrupted hub) is rerouted via a backup facility. $FFCB_{kl}$ accounts for the transportation cost of rerouting the flow through backup hub l when the first hub k is disrupted. $FFCB_{kl}$ is expressed as follows:

$$FFCB_{kl} = \sum_{i \neq k} \sum_{m \neq k} \sum_{j \neq k} C_{ilmj} x_{ikmj} + \sum_{i \neq k} \sum_{j \neq k} C_{illj}\, x_{ikkj} \tag{5}$$

The first term of the above expression calculates the rerouting cost where the path from origin i to destination j includes two hubs; the second term computes the cost where there is only one hub in the path from i to j.

$BFCB_{ml}$ deals with the cost of the rerouting part of the flow (via backup hub l) for which the disrupted facility is the second hub in the path from origin i to destination j. The expression for $BFCB_{ml}$ is as follows:

$$BFCB_{ml} = \sum_i \sum_{k \neq m} \sum_{j \neq m} C_{iklj} x_{ikmj} \tag{6}$$

In Eqs. (5) and (6) the purpose of imposing restrictions on index i and j is to exclude the transportation cost of the flow that initiate and/or end up at hub facilities. The cost of transporting this type of flow is dealt with separately in the last term of the following objective function. The limitation on index k and m prevents double-counting the cost of rerouting where the path from i to j include only one hub facility.

2.2. Initial mathematical formulation

We formulate the problem as a bi-objective optimization problem where the objectives are as follows:

$$F1 = \sum_i \sum_k \sum_m \sum_j C_{ikmj} x_{ikmj} \tag{7}$$

$$\begin{aligned} F2 = \sum_l &\left(\left(\sum_k FFC_k z_{ll} \right) - (FFC_l + BFC_l) \right) q_l \\ &+ \sum_k \sum_l FFCB_{kl} u_{kl} q_k \\ &+ \sum_m \sum_l BFCB_{ml} u_{ml} q_m \\ &+ \sum_i \sum_j \varphi_{ij} \lambda_{ij} (q_i z_{ii} + q_j z_{jj}) \end{aligned} \tag{8}$$

Objective *F1* computes the transportation cost in a regular situation while objective *F2* calculates the expected transportation cost (ETC) resulted from hub failures. More specifically, the first term in objective *F2* computes the transportation cost of the flow in part of the network(s) that is not affected by a disruption. This is achieved by subtracting the transportation cost of the flow that (initially) uses the disrupted hub (either as the first or the second hub) from the (total) transportation cost of the network in a regular situation. The second term accounts for the transportation cost of rerouting the flow through the backup hub when the first hub in the path from origin i to destination j is disrupted; the third term calculates the rerouting cost of the flow for which the disrupted facility is its second hub in the path from i to j. As mentioned earlier, it is assumed that when a hub is disrupted, it cannot send or receive any flow to or from other nodes in the network. Therefore, the last term in the objective *F2* (i.e., the forth term) penalizes the loss of flow/demand in disrupted situations where the source or destination of the flow is a hub. In our study, the penalty cost of losing a unit flow, φ_{ij}, is considered twice as much as the transportation cost of a unit flow between origin i and destination j.

The proposed model minimizes a weighted sum of the two objectives, $wF1 + (1-w)F2$ where $0 \le w \le 1$. The non-linear formulation of the single allocation p-hub median problem with hub unavailability (SApHM-HU) is presented as follows:

$$\text{Minimize } \; wF1 + (1-w)F2 \tag{9}$$

Subject to:

$$\sum_k z_{ik} = 1 \quad \forall i \tag{10}$$

$$\sum_k z_{kk} = p \tag{11}$$

$$z_{ik} \le z_{kk} \quad \forall i, k \tag{12}$$

$$\sum_m x_{ikmj} = z_{ik} \quad \forall i, j, k \tag{13}$$

$$\sum_k x_{ikmj} = z_{jm} \quad \forall i, j, m \tag{14}$$

$$\sum_{l \ne k} u_{kl} = z_{kk} \quad \forall k \tag{15}$$

$$u_{kl} \le z_{ll} \quad \forall k, l; \;\; k \ne l \tag{16}$$

$$FFC_k = \sum_i \sum_m \sum_j C_{ikmj} x_{ikmj} \quad \forall k \tag{17}$$

$$BFC_m = \sum_i \sum_{k \ne m} \sum_j C_{ikmj} x_{ikmj} \quad \forall m \tag{18}$$

$$FFCB_{kl} = \sum_{i \ne k} \sum_{m \ne k} \sum_{j \ne k} C_{ilmj} x_{ikmj} + \sum_{i \ne k} \sum_{j \ne k} C_{illj} x_{ikkj} \quad \forall k, l \tag{19}$$

$$BFCB_{ml} = \sum_i \sum_{k \ne m} \sum_{j \ne m} C_{iklj} x_{ikmj} \quad \forall m, l \tag{20}$$

$$x_{ikmj}, z_{ik}, u_{kl} \in \{0, 1\} \quad \forall i, j, k, m, l \tag{21}$$

$$FFC_k, \; BFC_m, \; FFCB_{kl}, \; BFCB_{ml} \ge 0 \quad \forall k, m, l \tag{22}$$

Constraints (10)–(14) are the classical constraints for the single allocation p-hub median problem [36]. Constraint (10) ensures every node is assigned to only one hub. Constraint (11) limits the number of hubs to be opened to a given number "p". Constraint (12) guarantees a node is assigned to exactly one hub. Constraints (13) and (14) ensure that all the traffic between an origin–destination pair has been routed via the hub sub-network.

Constraint (15) guarantees the disrupted node is a "hub" and it has only one backup. Constraint (16) ensures the backup node is a "hub" and it differs from the disrupted hub. Constraint (17) calculates the total transportation cost for part of the flow that utilizes hub k as the first hub; constraint (18) calculates the total transportation cost for a part of the flow that utilizes hub m as the second hub. Constraint (19) accounts for the rerouting cost of the flow from origin i to destination j via backup l when the first hub k is disrupted; constraint (20) deals with the rerouting cost when the second hub m is disrupted. Constraints (21) and (22) are standard integrality constraints.

The proposed model for SApHM-HU described above is a mixed integer quadratic program (*MIQP*). The resulting formulation for SApHM-HU has $2n + 4n^2 + n^4$ variables and $1 + 4n + 4n^2 + 2n^3$ linear constraints. In the next sub-section, we present a mixed integer linear formulation (MILP) for the problem and show that a standard linearization will not significantly increase the problem's complexity with respect to the additional number of variables and/or constraints.

2.3. Linearization and models complexity comparison

The non-linear terms in the above objective function resulted from the multiplication of binary and non-binary variables of u, z, FFC_k, $FFCB_{kl}$ and $BFCB_{ml}$. To linearize the model, these terms are substituted by three continuous variables ψ_{kl}, Γ_{kl} and ξ_{ml} as follows:

$$\psi_{kl} = FFC_k z_{ll} \tag{23}$$

$$\Gamma_{kl} = FFCB_{kl} \, u_{kl} \tag{24}$$

$$\xi_{ml} = BFCB_{ml} u_{ml} \tag{25}$$

To enforce the above equations, constraints (27)–(39) are added to the formulations.

The resulting linear model is then presented as follows:

$$\text{Minimize } \; wF1 + (1-w)F2 \tag{26}$$

Subject to: Eqs. (10)–(20)

$$\psi_{kl} \le v_1 \times z_{ll} \quad \forall k, l \tag{27}$$

$$\psi_{kl} \le FFC_k \quad \forall k, l \tag{28}$$

$$\psi_{kl} \ge FFC_k - v_1(1 - z_{ll}) \quad \forall k, l \tag{29}$$

$$\Gamma_{kl} \le v_2 \times u_{kl} \quad \forall k, l \tag{30}$$

$$\Gamma_{kl} \le FFCB_{kl} \quad \forall k, l \tag{31}$$

Table 2
Models complexity comparison.

Problem	Number of variables	Additional variables	Number of constraints	Additional constraints
SApHMP[a]	n^2+n^4	–	$1+n+n^2+2n^3$	–
SApHM-HU (MIQP)	$2n+4n^2+n^4$	$2n+3n^2$	$1+4n+4n^2+2n^3$	$3n+3n^2$
SApHM-HU (MILP)	$2n+7n^2+n^4$	$2n+6n^2$	$1+4n+13n^2+2n^3$	$3n+12n^2$

[a] The classical p-hub location problem formulation proposed by Skorin-Kapov et al. [35].

$$\Gamma_{kl} \ge FFCB_{kl} - v_2(1-u_{kl}) \quad \forall k,l \tag{32}$$

$$\xi_{ml} \le v_3 \times u_{ml} \quad \forall m,l \tag{33}$$

$$\xi_{ml} \le BFCB_{ml} \quad \forall m,l \tag{34}$$

$$\xi_{ml} \ge BFCB_{ml} - v_3(1-u_{ml}) \quad \forall m,l \tag{35}$$

$$0 \le \psi_{kl} \le v_1 \quad \forall k,l \tag{36}$$

$$0 \le \Gamma_{kl} \le v_2 \quad \forall k,l \tag{37}$$

$$0 \le \xi_{ml} \le v_3 \quad \forall m,l \tag{38}$$

$$\psi_{kl}, \Gamma_{kl}, \xi_{ml} \ge 0 \quad \forall k,l,m,l \tag{39}$$

$$x_{ikmj}, z_{ik}, u_{kl} \in \{0,1\} \quad \forall i,j,k,m,l \tag{40}$$

$$FFC_k, BFC_m, FFCB_{kl}, BFCB_{ml} \ge 0 \quad \forall k,m,l \tag{41}$$

where

$$F1 = \sum_i \sum_k \sum_m \sum_j C_{ikmj} x_{ikmj} \tag{42}$$

and

$$\begin{aligned} F2 = &\sum_l \left(\sum_k \psi_{kl} - (FFC_l + BFC_l) \right) q_l \\ &+ \sum_k \sum_l \Gamma_{kl} q_k + \sum_m \sum_l \xi_{ml} q_m \\ &+ \sum_i \sum_j \varphi_{ij} \lambda_{ij} (q_i z_{ii} + q_j z_{jj}) \end{aligned} \tag{43}$$

The above linear model has $2n+7n^2+n^4$ variables and $1+4n+13n^2+2n^3$ constraints. A comparison of the linear and non-linear models for SApHM-HU shows that the standard linearization does not significantly increase the number of variables (see Table 2). In other words, both mixed integer quadratic and mixed integer linear programming formulations have to deal with $O(n^4)$ and $O(n^3)$ variables and constraints respectively. The total number of variables and constraints in the two proposed models is presented in Table 2 alongside the classical single allocation p-hub median problem (SApHMP) formulation developed by Skorin-Kapov et al. [36].

Despite a moderate increase in the number of variables and constraints, it is still difficult to solve the linear version of SApHM-HU by exact methods as will be shown in the computational result section. Therefore, the best way forward would be to design a meta-heuristic for this purpose. In our study, we opted for an evolutionary approach namely a genetic algorithm which is described next.

3. A genetic-based algorithm

Genetic algorithms (GAs) simulate the evolution process of species reproduction [17]. Unlike other meta-heuristics such as simulated annealing [20] and tabu search [16] that work with a single solution, GA deals with a population of solutions. Various types of genetic algorithms have been successfully applied to optimization problems including location-allocation, assignment, etc. Examples of such studies are Cunha and Silva [11], Topcuoglu et al. [38], Kratica et al. [23] and Helm [1]. For a comprehensive study of hub-and-spoke problems and solution techniques the reader is referred to Alumur and Kara [2].

In this study, we present a GA-based algorithm to solve the SApHM-HU. The main characteristics of the proposed method are the introduction of an efficient solution representation and a crossover operator that jointly aim to maintain solution feasibility during the reproduction phase.

The flowchart of the proposed GA is given in Fig. 2. In the following subsections, we elaborate on the way we implement our proposed GA focusing on the solution representation, the choice of the crossover and mutation operators, and finally the backup selection procedure.

3.1. Solution representation (hub-based approach)

In the context of hub-and-spoke problem, the commonly used chromosome representation is the binary representation in which solutions are encoded as strings of zeros (0) and ones (1) [11,38].

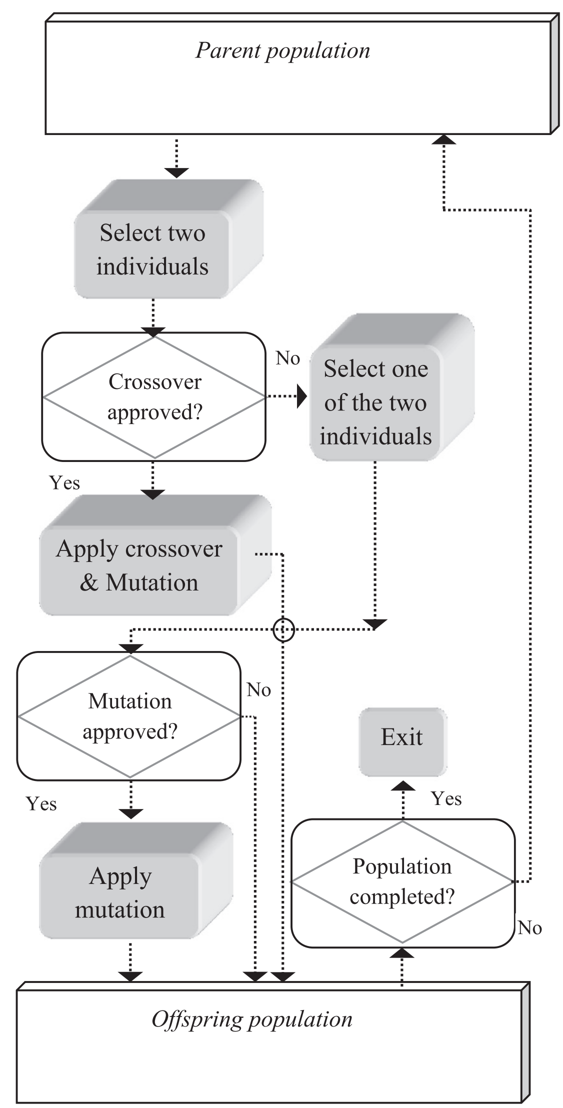


Fig. 2. Flowchart of the proposed genetic algorithm.

However, a chromosome represented this way may yield an infeasible solution if it is subjected to even a minor perturbation. To alleviate this deficiency, a repair mechanism is often introduced to transform such an infeasible solution into a feasible one. This additional requirement is one of the key factors for a low performance including the excessive computational time associated with some GAs reported in the literature. In this study, we propose an efficient solution representation and crossover operator that jointly aim to maintain solution feasibility during the reproduction phase.

We term our solution representation scheme a *hub-based* approach. This uses an array (or string) with the length of $1 \times n$ where the length of the array, n, corresponds to the total number of nodes in the network. Decoding the string from left to right, the first location corresponds to node number 1, the second location to node number 2, and so on. Each location on the string (i.e., a gene) contains a number which may or may not be the same as the "column number". These numbers represent the hub facilities in the network to which one or more nodes (i.e., column number) are allocated. Each hub node is allocated to itself; this is shown where a column number matches a hub number. Fig. 3 illustrate a typical solution to the problem with 10 nodes and three hubs. Nodes 5, 7, and 9 are assumed to be hub facilities. The three nodes are allocated to themselves but the other nodes in the network are assigned to one of these hubs.

3.2. Crossover operator

Given the solution representation described above, we present a special type of crossover operator to generate new offspring solutions. The proposed crossover operator ensures feasibility of the resulting solutions and its main steps are given in Fig. 4.

In transferring genes from each of the parent chromosomes, the following cases may arise:

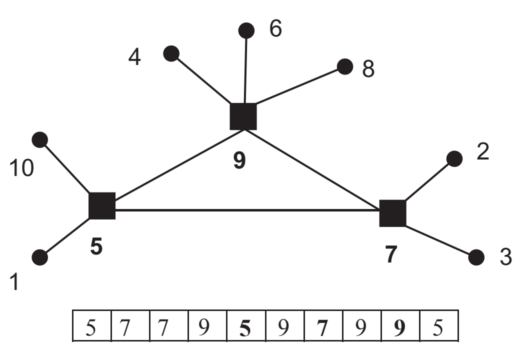


Fig. 3. Solution representation scheme.

Step 1: Construct a template array with the length of *n*

Step 2: Select two parent chromosomes form the current population

Step 3: Choose one of the two parents randomly (e.g., parent 1)

DO

Step 4: Scan the parent chromosome in hand and transfer the first available gene (i.e., a hub) and its counterparts (i.e., all nodes that are assigned to that particular hub) to the offspring chromosome

Step 5: Select the other parent

UNTIL the offspring chromosome is completely constructed or no further gene could be transferred to the offspring chromosome from any of the two parents

Step 6: If the offspring chromosome is completed, **Stop**; otherwise ignore the current offspring and **GO** to **Step 1**

Fig. 4. A pseudo-code for the proposed crossover operator.

Case 1: The corresponding location of the selected gene in the offspring chromosome has been already occupied by another gene. In this case, the selected gene is discarded and the donor chromosome is further scanned to find another candidate gene to transfer.

Case 2: All transferable genes in two parent chromosomes have been transferred to the offspring chromosome but the new solution is not being completely constructed. In other words, there are one or more locations (nodes) in the offspring chromosome for which no hub centre have been decided. In this case, if the offspring chromosome contains p hubs then the empty locations in the offspring chromosome are filled by randomly allocating the undecided nodes (nodes with no assigned values) to the existing hub centres. Otherwise, the solution is discarded and another offspring is generated using a new template array and a new pair of parents.

3.2.1. An illustrative example of our crossover operation

An illustrative example of a crossover operation is presented in Fig. 5. In parent1 chromosome, the hub centres are node 1, 4, and 7 while in parent2 they are assumed to be nodes 3, 5 and 7. Parent1 is selected to donate the first gene. The crossover operation begins with constructing an empty array with the length of 10. Parent1 chromosome is then scanned from left to right and the first gene (i.e., the first hub centre) is selected. The first selected gene is "1" which is placed in the "first" position (from the left side) of the offspring chromosome. Next the other genes with the same value of "1" (i.e., second and third genes) are transferred to their corresponding places in the offspring chromosome. Once the first gene is successfully transferred from parent1 into the offspring chromosome, the second parent (i.e., parent2) is scanned from left to right to search for a candidate gene that could be transferred into the offspring chromosome. In this case, the first candidate to be examined is hub centre "3". The gene with value "3" should be placed in the third position of the offspring chromosome. However, as the third column in the offspring chromosome has been already occupied by hub centre "1", this gene is discarded and values of all genes that contain the same value, "3", in parent2 are set to zero. As the gene with value "3" could not be transferred to the offspring's template chromosome, parent2 chromosome is scanned again to search for a new transferable gene; the next candidate is gene "5". The fifth position in the offspring chromosome is free then hub centre "5" (gene with value 5) could be placed in the fifth location of the offspring chromosome. Subject to availability of space in the offspring chromosome, other genes with value "5" in parent2 are next transferred to their corresponding locations in the offspring chromosome. The offspring chromosome has not been completed yet; therefore, parent1 is re-examined to find another gene that could be transferred into the offspring chromosome. Similar to gene "3" of parent 2, gene "4" in parent 1 chromosome could not be transferred as its corresponding location in the offspring chromosome has been occupied by another gene i.e., gene "5". Copying the next transferable gene from parent1 chromosome (i.e., gene "7") to the offspring chromosome ends the crossover

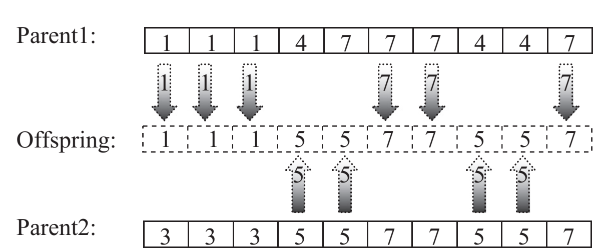


Fig. 5. Example of crossover operation.

operation. It is worthwhile to mention that selecting each one of the two parents as the first donor at the beginning of the crossover operation will likely lead to a structurally different offspring chromosome.

3.3. Mutation operator

In our GA, the mutation operation is performed by randomly selecting two unidentical genes that represent non-hub nodes in the network and swapping their positions. The selected genes must be unidentical to insure population diversity and they have to be non-hub to avoid generating infeasible solutions.

3.4. Backup Hub selection

The objective of SApHM-HU is to select the location of the hubs, allocate non-hub nodes to the hub facilities, and determine a backup facility for each hub in the network. The first and the second decisions are explicitly modeled in the solution representation described above. To decide on the backup facilities, we examined the following two approaches:

(i) In the first method, once a new solution is generated a backup facility is selected randomly for every hub in its corresponding network.
(ii) In the second approach, upon the generation of a new solution, $p-1$ candidate solutions are generated by assigning one of the existing hubs as the backup facility for the first hub in the network. Note that for each hub in the network there exist $p-1$ candidate backup facilities. The total cost associated with each candidate solution is then calculated and the one with lowest cost is selected.

To further clarify the second backup selection approach, consider a typical solution for the previous problem in Fig. 6. Here if hub "5" is disrupted, either hub "7" or hub "9" could be used as the backup hub for the assumed disrupted hub "5". To decide on the backup, the transportation costs of the two available options are calculated and compared. The backup node with the lower cost is selected. This process continues to select a backup for the other two hubs in the network (i.e., hubs "7" and hub "9").

In this study, we performed a small experiment to evaluate and compare the performance of the above two strategies in selecting backup hubs. The results of our experiment for the problem instances with up to five hubs show that the second approach provides better solutions without significantly increasing the computational times. However, as the number of hubs increases (e.g., greater than 5) the first approach which randomly selects backup hubs is expected to be computationally more efficient as the second one enumerates all the existing hubs (i.e., $p-1$).

4. Computational results

4.1. Experimental design

We tested our algorithm on 144 benchmark problems with 10, 15, 20, 25, 55 and 81 nodes derived from U.S. Civil Aeronautics Board (CAB) [32] and Turkish Postal System (TR) datasets [12]. The problem instances are generated by setting the number of nodes n to 10, 15, 20, 25 55 and 81; the number of hubs p to be opened to 3 and 5; the discount factor α to 0.2, 0.4 and 0.8; and the objective weight w to 0.3, 0.5, and 0.7. The coefficients of collection χ and distribution cost δ are set to 1 per unit for all test problems ($\chi=\delta=1$). The probability of hub failures, q_i, for all nodes in the network $i=1, \ldots, n$ is generated randomly and selected from uniform distribution [0.1, 0.3].

The GA parameters are decided as follows. The population size is set to 150 for all test problems; the probabilities of crossover and mutation are set to 0.85 and 0.9 respectively. The computational time for problem with 10 and 15 nodes is set to 10 and 20 s respectively; for problems with 20 nodes and three hubs it is set to 20 s; for 20 nodes and five hubs it is 30 s; for problems with 25 nodes and three hubs it is set to 40 s; for 25 nodes and five hubs the computational times is set 60 s; for 55 nodes (three and five hubs) it is set to 90 and 120 s; for 81 nodes it is set to 150 s. Each test problem is run 20 times and the best results are reported. The MIP version of SApHM-HU model is coded in AIMMS and solved using CPLEX 12.4 with CPLEX options set to their default values. The algorithms are run on a Dell Intel Core PC with 2.40 GHz processor with 2 GB of RAM.

4.2. Comparison vs. optimal solution (small instances)

The computational performance of the proposed GA and CPLEX is presented in Table 3. In Table 3, we present the computational results for 36 problems with 10 nodes and three and five hubs derived from both CAB [32] and TR [12]datasets. In this table, we report the Total Cost (TC) which sums the Regular and the Expected Transportation Costs (RTC and ETC), location of the hubs, backup hubs, and the processing time to obtain the optimal/best solutions. For instance, in the first benchmark problem where $n=10$, $p=3$, $\alpha=0.2$, $w=0.3$ the total cost of the optimal solution is 367.24 units. The hub facilities to be opened are located at nodes 4, 2, and 7 and the backup hubs are 2, 4, and 4 respectively (hub 4 serves as the backup up for both hub 2 and hub 7). We use the same parameters and input data including the inter-hub discount factors (α), objective weights (w), coefficients of collection and distribution (χ and δ) and probability of hub failures (q_i) in both GA and CPLEX.

Using CPLEX we solved all 36 benchmark problems to optimality. However, our results show that the algorithm requires significantly longer computing time when compared against that of the GA. The average computing times returned by CPLEX and

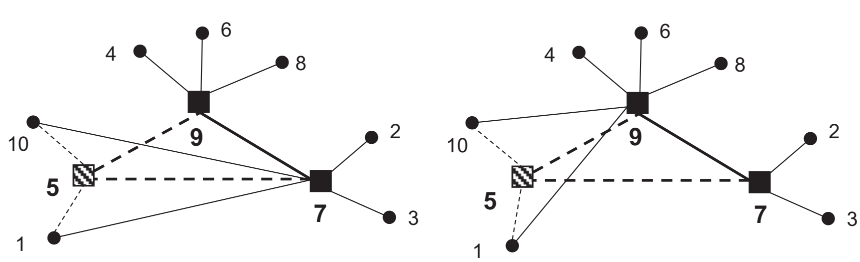


Fig. 6. Backup selection.

Table 3
Computational results for the GA and CPLEX.

n[a]	p[b]	α[c]	w[d]	GA best				CPLEX		%GAP
				Hub locations	Backup hubs	TC[e]	Time (s)	TC[e]	Time (s)	
10 (CAB)	3	0.2	0.3	4, 2, 7	2, 4, 4	367.24	0.98	367.24	6160	0.00
			0.5	4, 2, 7	2, 4, 4	409.89	0.45	409.89	460	0.00
			0.7	4, 2, 7	2, 4, 4	452.54	0.64	452.54	434	0.00
		0.4	0.3	4, 2, 7	2, 4, 4	401.54	0.72	401.54	3140	0.00
			0.5	4, 2, 7	2, 4, 4	457.13	0.97	457.13	714	0.00
			0.7	4, 2, 7	2, 4, 4	512.72	1.20	512.72	336	0.00
		0.8	0.3	4, 2, 7	2, 4, 4	470.15	1.30	470.15	3700	0.00
			0.5	4, 2, 7	2, 4, 4	551.62	0.62	551.62	862	0.00
			0.7	4, 9, 7	9, 4, 4	621.87	1.40	621.87	329	0.00
	5	0.2	0.3	1, 2, 4, 7, 8	8, 4, 2, 1, 7	375.46	4.20	375.46	1970	0.00
			0.5	1, 2, 4, 7, 8	8, 4, 2, 1, 7	369.16	6.10	369.03	322	0.00
			0.7	1, 2, 4, 7, 8	8, 4, 2, 1, 4	362.78	3.60	362.78	401	0.00
		0.4	0.3	1, 2, 4, 7, 8	8, 4, 2, 1, 7	432.36	3.20	432.53	2316	0.00
			0.5	1, 2, 4, 7, 8	8, 4, 2, 1, 2	437.50	3.00	437.50	373	0.00
			0.7	1, 2, 4, 9, 7	7, 9, 7, 4, 4	443.94	7.60	442.46	276	0.33
		0.8	0.3	1, 2, 4, 7, 8	7, 4, 2, 1, 2	547.03	2.20	547.03	2924	0.00
			0.5	1, 2, 4, 7, 8	2, 4, 2, 1, 2	574.43	5.90	574.43	832	0.00
			0.7	1, 9, 4, 7, 8	8, 4, 8, 1, 7	588.80	2.80	588.85	684	0.00
10 (TR)	3	0.2	0.3	1, 2, 7	2, 1, 1	293.24	5.84	293.24	925	0.00
			0.5	1, 2, 7	2, 1, 1	316.79	0.30	316.79	253	0.00
			0.7	1, 2, 7	2, 1, 1	340.35	0.47	340.34	178	0.00
		0.4	0.3	1, 2, 7	2, 1, 1	331.74	0.50	331.74	1436	0.00
			0.5	1, 2, 7	2, 1, 1	371.96	0.30	331.74	241	0.00
			0.7	1, 2, 7	2, 1, 1	412.18	0.56	412.18	208	0.00
		0.8	0.3	1, 2, 7	2, 1, 2	405.72	0.27	405.72	2660	0.00
			0.5	1, 2, 7	2, 1, 1	475.58	0.42	475.58	350	0.00
			0.7	1, 2, 7	2, 1, 2	543.38	0.61	543.38	206	0.00
	5	0.2	0.3	1, 2, 4, 7, 8	8, 8, 7, 8, 2	333.59	1.02	333.59	867	0.00
			0.5	1, 2, 4, 7, 8	8, 8, 7, 8, 2	324.55	1.06	324.55	146	0.00
			0.7	1, 2, 4, 6, 7	2, 6, 7, 2, 6	288.63	1.37	288.63	97	0.00
		0.4	0.3	1, 2, 4, 7, 8	8, 8, 1, 8, 2	393.88	0.46	393.88	712	0.00
			0.5	1, 2, 4, 7, 8	8, 8, 1, 8, 2	395.74	3.18	395.74	248	0.00
			0.7	1, 2, 4, 6, 7	2, 6, 1, 1, 6	391.22	2.61	391.22	151	0.00
		0.8	0.3	1, 2, 4, 7, 8	8, 1, 1, 8, 1	513.05	0.90	513.05	1516	0.00
			0.5	1, 2, 4, 7, 8	8, 1, 1, 8, 1	537.11	2.50	537.11	382	0.00
			0.7	1, 2, 4, 7, 8	8, 1, 1, 8, 1	561.18	1.64	561.18	288	0.00

[a] Number of nodes.
[b] Number of hubs.
[c] Discount factor.
[d] Objective weight.
[e] Total cost.

the GA over all 36 instances presented in Table 3 is 1030.47 and 1.97 s respectively. The GA solved 35 (out of 36) instances to optimality in a very short computing time while a near optimal solution was located within a 0.3% duality gap only.

4.3. The GA performance (large instances)

Tables 4 and 5 present the computational performance of the GA on problem instances with 15, 20, 25, 55 and 81 nodes which are taken from CAB and TR datasets. The average times to obtain the best solutions for problems with 15, 20, and 25 nodes (CAB instances) are 10.7, 18.1, and 39.8 s respectively; for problems with 25 (TR instances), 55 and 81 nodes are 30.9, 97.3 and 139.7 s respectively. The average computational time to obtain the best solutions over all 108 instances is 56.1 s. For relatively larger problem instances i.e., problems with 15, 20, 25, 55 and 81 nodes CPLEX fails to produce solutions due to excess of memory.

A further examination of the final solutions to the benchmark problems presented in Tables 3–5 shows that, especially for low values of the objective weight (e.g., $w=0.3$), the algorithm is more likely to locate hub facilities in areas with low probability of failure as long as the regular transportation cost (RTC) is not substantially high. For instance, for the problem with 15 nodes and three hubs ($\alpha=0.2$; $w=0.3$), the hub facilities are located at nodes 6 (with probability of failure=0.12), 10(0.10), and 15 (0.11). However, as the value of the objective weight increases to 0.7 ($w=0.7$) the new location of the hubs found to be at nodes 4(0.16), 10(0.10) and 12 (0.24). This is not surprising as a high probability of hub failures elevates the expected transportation cost which is the primary concern when the objective weight is low.

Fig. 7 illustrates approximate locations and names of the 25 US cities of CAB dataset. For the TR dataset, this information could be found in the Appendix.

4.4. A general discussion

To learn more about the pros and cons of considering hub unavailability at the design stage, we examine the two optimal solutions provided by SApHM-HU and SApHM models for the problem with 10 nodes and 3 hubs. Here, the inter-hub discount factor and the objective function weight are 0.2 and 0.3 respectively. As presented in Table 3, the optimal solution to the SApHM-HU problem suggests a network in which Chicago (nodes 4), Baltimore (node 2), and Dallas-Fw (node 7) are hubs (see Fig. 8). The (unweighted) regular and expected transportation costs of the network are 516.52 and 303.26 units respectively. With regard to

Table 4
Computational results for the GA: CAB dataset.

n[a]	p[b]	α[c]	w[d]	RTC[e]	ETC[f]	TC[g]	Time (s)	Hub locations	Backup hubs
15	3	0.2	0.3	318.70	298.07	616.77	13.9	6, 10, 15	15, 15, 10
			0.5	495.65	238.35	734.00	8.2	6, 4, 10	4, 6, 4
			0.7	583.33	207.88	791.21	10.8	4, 10, 12	10, 4, 10
		0.4	0.3	344.97	322.28	667.25	16.2	9, 10, 15	15, 15, 10
			0.5	538.93	248.81	787.74	11.6	6, 4, 10	4, 10, 4
			0.7	713.62	179.20	892.82	2.8	4, 6, 7	6, 4, 4
		0.8	0.3	373.96	377.98	751.94	2.6	4, 6, 10	6, 4, 4
			0.5	577.87	302.28	880.15	3.9	9, 4, 11	4, 11, 4
			0.7	806.52	176.46	982.98	2.3	6, 4, 11	6, 4, 6
	5	0.2	0.3	199.67	533.37	733.04	10.0	6, 4, 10, 15, 12	4, 10, 4, 4, 15
			0.5	283.18	409.06	692.24	16.6	6, 4, 10, 12, 14	4, 10, 4, 14, 12
			0.7	373.60	261.22	634.82	12.1	6, 4, 7, 12, 14	4, 7, 4, 7, 7
		0.4	0.3	289.72	516.61	806.33	18.0	6, 4, 11, 10, 15	4, 15, 10, 15, 10
			0.5	388.84	426.48	815.32	17.7	6, 4, 10, 11, 12	4, 11, 11, 4, 11
			0.7	489.61	293.90	783.51	11.2	6, 4, 7, 12, 14	4, 7, 4, 7, 4
		0.8	0.3	334.08	566.51	900.59	13.6	6, 4, 10, 11, 15	4, 11, 11, 15, 11
			0.5	556.77	407.91	964.68	11.1	6, 4, 11, 10, 15	4, 11, 10, 15, 10
			0.7	751.80	262.50	1014.30	9.5	5, 6, 4, 10, 11	6, 4, 11, 11, 4
20	3	0.2	0.3	247.04	281.97	529.01	22.2	4, 17, 10	17, 4, 4
			0.5	382.88	226.62	609.50	8.9	4, 17, 19	17, 4, 4
			0.7	507.18	162.64	669.82	7.2	4, 17, 12	17, 4, 4
		0.4	0.3	284.21	296.09	580.31	15.9	4, 17, 10	17, 4, 4
			0.5	466.49	210.99	677.49	8.7	4, 17, 10	17, 4, 4
			0.7	593.44	170.16	763.60	6.8	4, 17, 12	17, 4, 4
		0.8	0.3	332.00	333.13	665.12	26.6	1, 17, 4	4, 4, 1
			0.5	548.63	237.37	785.99	6.4	1, 17, 4	4, 4, 1
			0.7	773.36	141.82	915.18	8.6	1, 17, 4	17, 4, 1
	5	0.2	0.3	164.80	389.92	554.72	25.1	1, 17, 4, 7, 19	4, 4, 17, 19, 7
			0.5	288.34	267.14	555.47	23.8	1, 17, 4, 10, 19	4, 4, 1, 1, 10
			0.7	360.36	191.88	552.24	25.7	1, 17, 4, 7, 12	17, 4, 1, 4, 7
		0.4	0.3	210.70	414.85	625.55	20.8	1, 17, 4, 10, 19	4, 4, 1, 1, 10
			0.5	356.61	329.31	685.92	22.8	1, 17, 4, 20, 19	4, 20, 20, 4, 4
			0.7	483.55	231.95	715.50	25.6	1, 6, 17, 4, 12	4, 17, 6, 6, 4
		0.8	0.3	318.79	488.56	807.35	26.8	1, 17, 3, 4, 10	17, 4, 10, 1, 4
			0.5	536.79	351.46	888.24	20.1	1, 17, 3, 4, 10	4, 4, 10, 1, 4
			0.7	727.32	224.20	951.53	23.5	1, 17, 3, 4, 7	17, 4, 7, 1, 4
25	3	0.2	0.3	262.85	305.82	568.67	39.5	5, 25, 19	25, 5, 5
			0.5	447.82	219.84	667.66	42.1	5, 25, 19	25, 5, 5
			0.7	581.29	160.90	742.19	26.7	5, 25, 12	25, 5, 5
		0.4	0.3	316.12	301.25	617.38	23.2	5, 25, 8	25, 5, 5
			0.5	511.21	214.87	726.09	15.3	5, 25, 8	25, 5, 5
			0.7	662.44	168.46	830.90	31.4	5, 25, 12	25, 5, 5
		0.8	0.3	365.17	328.24	693.40	32.7	5, 25, 8	25, 5, 5
			0.5	628.86	238.94	867.80	29.2	5, 25, 8	25, 5, 5
			0.7	862.89	150.93	1013.82	31.9	5, 20, 8	20, 5, 5
	5	0.2	0.3	239.60	391.36	630.96	55.2	25, 5, 3, 10, 19	5, 25, 25, 5, 10
			0.5	399.23	299.10	698.33	54.3	5, 25, 3, 21, 19	21, 5, 25, 5, 21
			0.7	431.13	246.87	678.00	59.1	25, 17, 4, 10, 12	17, 25, 25, 4, 10
		0.4	0.3	289.11	435.57	724.68	35.3	25, 5, 10, 8, 20	20, 25, 5, 10, 5
			0.5	483.24	307.87	791.11	53.6	5, 25, 20, 10, 8	20, 20, 5, 5, 10
			0.7	595.26	248.37	843.63	50.4	5, 17, 10, 12, 25	25, 25, 5, 10, 17
		0.8	0.3	359.10	509.39	868.49	47.9	25, 20, 5, 10, 8	20, 5, 20, 5, 10
			0.5	571.44	393.21	964.65	39.1	21, 25, 5, 20, 8	5, 21, 21, 21, 21
			0.7	794.99	259.60	1054.59	49.1	5, 25, 4, 20, 8	20, 20, 5, 5, 20

[a] Number of nodes.
[b] Number of hubs.
[c] Discount factor.
[d] Objective weight.
[e] Regular transportation cost.
[f] Expected transportation cost.
[g] Total cost.

backup facilities, a solution suggests hub 2 (i.e., Baltimore) as the backup for hub 4 (Chicago) and hub 4 (Chicago) as the backup for facilities 2 (Baltimore) and 7 (Dallas-Fw).

The optimal solution to the same problem where the objective is merely to determine the ideal hub locations and the allocation of non-hubs (i.e., the optimal solution to the SApHM problem) is depicted in Fig. 1. The regular transportation cost of this network is 491.93 [36]. To calculate the expected transportation cost of the network presented in Fig. 1, we simulate the situations in which one hub at a time becomes unavailable and the most economically attractive hub in the network is employed as the backup facility. The expected transportation cost is then calculated by measuring the routing and rerouting as well as the demand loss costs. For a fair comparison, the same probability of hub failures are utilized to calculate the expected transportation costs of the two solutions provided by SApHM-HU and SApHM models.

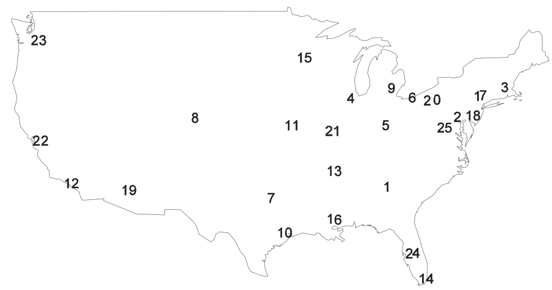
Table 5
Computational results for the GA: TR dataset.

n[a]	p[b]	α[c]	w[d]	RTC[e]	ETC[f]	TC[g]	Time (s)	Hub locations	Backup hubs
25	3	0.2	0.3	205.12	214.00	419.13	13.9	14, 16, 17	16, 17, 16
			0.5	270.83	196.73	467.56	13.3	19, 16, 15	16, 15, 16
			0.7	382.27	124.26	506.53	9.8	11, 3, 15	3, 11, 3
		0.4	0.3	224.33	224.37	448.71	27.0	14, 16, 17	16, 17, 16
			0.5	340.73	186.90	527.63	11.8	14, 16, 15	16.15.16
			0.7	437.73	133.75	571.48	23.1	11, 3, 15	3, 15, 3
		0.8	0.3	256.49	239.95	496.43	6.6	16, 17, 13	17, 16, 17
			0.5	435.51	172.65	608.16	12.2	16, 17, 13	17, 16, 17
			0.7	598.47	102.83	701.30	12.9	16, 17, 13	17, 16, 17
	5	0.2	0.3	133.16	304.95	438.11	37.4	19, 16, 17, 15, 13	16, 17, 15, 17, 17
			0.5	205.71	238.39	444.10	48.1	19, 3, 16, 13, 15	16, 15, 3, 15, 3
			0.7	274.31	140.41	414.72	36.5	19, 3, 16, 13, 15	16, 15, 3, 15, 3
		0.4	0.3	178.66	325.31	503.97	51.0	14, 16, 13, 15, 17	16, 17, 17, 17, 15
			0.5	258.59	262.76	521.35	57.2	11, 3, 16, 15, 13	16, 15, 3, 3, 15
			0.7	353.54	162.29	515.83	41.1	19, 3, 17, 15, 13	3, 17, 13, 17, 17
		0.8	0.3	233.02	396.66	629.68	51.2	16, 14, 17, 15, 13	14, 16, 15, 17, 17
			0.5	359.14	307.15	666.29	57.1	16, 14, 3, 15, 13	3, 16, 15, 3, 15
			0.7	472.99	211.39	684.38	45.8	1, 3, 16, 15, 13	16, 16, 3, 3, 15
55	3	0.2	0.3	205.13	223.78	428.91	71.4	55, 18, 26	18, 26, 18
			0.5	325.92	159.68	485.60	81.1	55, 18, 26	18, 26, 18
			0.7	469.51	109.17	578.68	85.4	19, 18, 26	18, 26, 18
		0.4	0.3	240.28	222.15	462.43	85.9	55, 14, 18	14, 18, 14
			0.5	368.81	168.52	537.33	89.4	55, 18, 26	18, 26, 18
			0.7	554.71	123.14	677.85	86.6	55, 45, 26	45, 26, 45
		0.8	0.3	284.93	252.30	537.23	81.9	29, 18, 14	14, 29, 29
			0.5	468.18	176.17	644.35	87.1	55, 18, 14	14, 14, 18
			0.7	645.15	116.11	761.26	69.2	29, 18, 26	18, 26, 18
	5	0.2	0.3	191.80	313.57	505.37	113.8	55, 14, 7, 33, 26	14, 26, 26, 7, 33
			0.5	306.39	214.93	521.32	97.2	55, 18, 14, 26, 33	14, 26, 55, 33, 18
			0.7	434.87	129.28	564.15	111.2	55, 14, 18, 33, 26	14, 18, 26, 18, 18
		0.4	0.3	222.74	336.55	559.29	113.7	25, 55, 18, 26, 14	55, 25, 26, 18, 18
			0.5	360.81	250.00	610.81	114.4	34, 29, 18, 26, 14	29, 14, 29, 18, 18
			0.7	497.35	173.84	671.20	112.0	55, 2, 31, 32, 26	2, 55, 32, 26, 32
		0.8	0.3	278.56	404.40	682.96	117.3	25, 34, 9, 18, 14	34, 25, 14, 14, 18
			0.5	464.34	285.35	749.69	119.2	25, 18, 14, 9, 26	14, 26, 18, 14, 18
			0.7	649.52	181.94	831.46	113.8	29, 7, 18, 26, 33	18, 26, 26, 29, 7
81	3	0.2	0.3	257.17	275.59	532.75	137.7	42, 44, 14	14, 42, 42
			0.5	426.98	210.62	637.60	148.2	51, 44, 6	44, 51, 51
			0.7	567.18	142.81	709.99	143.9	51, 16, 6	6, 6, 51
		0.4	0.3	277.83	289.78	567.61	141.5	51, 6, 9	6, 51, 6
			0.5	452.04	219.60	671.64	142.4	42, 44, 6	6, 6, 42
			0.7	630.38	132.19	762.57	149.5	50, 6, 9	6, 50, 6
		0.8	0.3	311.10	287.14	598.25	118.0	51, 71, 4	71, 51, 71
			0.5	506.39	215.59	721.99	141.4	51, 6, 4	6, 51, 51
			0.7	710.11	129.42	839.53	125.2	51, 6, 4	6, 51, 51
	5	0.2	0.3	278.43	409.11	687.54	150.0	42, 14, 47, 5, 9	5, 42, 5, 42, 42
			0.5	460.41	285.55	745.96	145.5	71, 4, 42, 9, 47	42, 47, 9, 42, 71
			0.7	555.43	215.57	771.00	124.9	58, 6, 18, 34, 45	6, 18, 6, 6, 6
		0.4	0.3	290.47	446.37	736.84	133.7	71, 14, 4, 5, 69	14, 71, 69, 71, 5
			0.5	448.83	338.47	787.30	150.1	50, 71, 4, 9, 34	71, 50, 71, 71, 71
			0.7	630.70	214.54	845.24	148.2	71, 51, 6, 9, 34	6, 71, 71, 6, 6
		0.8	0.3	316.53	454.26	770.79	122.9	71, 42, 4, 5, 9	42, 71, 5, 71, 71
			0.5	512.14	329.28	841.42	150.0	6, 4, 5, 9, 51	51, 51, 6, 6, 6
			0.7	717.64	201.67	919.32	142.0	6, 4, 5, 9, 71	71, 71, 71, 6, 6

[a] Number of nodes.
[b] Number of hubs.
[c] Discount factor.
[d] Objective weight.
[e] Regular transportation cost.
[f] Expected transportation cost.
[g] Total cost.

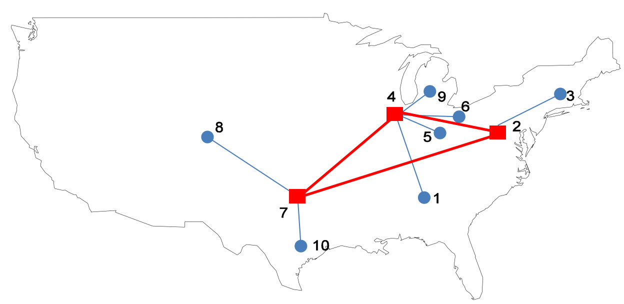
Our simulation results indicate that in the network without hub unavailability, the consideration of the minimum expected transportation cost incurs when hub 4 (Chicago) serves as the backup for hubs 6 (Cleveland) and 7 (Dallas-Fw), and hub 7 (Dallas-Fw) supports hub 4 (Chicago). The expected transportation cost of the network is then 403.32 units.

The above analysis shows that the regular transportation cost of the solution provided by SApHM-HU (516.52 unit) model is 4.8% higher than that of the solution provided by the model without hub unavailability consideration (491.93 units). Its expected transportation cost however, is in fact 25% lower (303.26 vs. 403.32). Managers may hesitate to undertake such a large increase in the regular transportation cost but they may be willing to spend 4.8% more in order to reduce the expected transportation cost which may result from possible hub failures. This accounts for about 25% as shown in the above problem instance.



1 Atlanta	6 Cleveland	11 Kansas City	16 New Orleans	21 St. Louis
2 Baltimore	7 Dallas-Fw	12 Los Angeles	17 New York	22 San Francisco
3 Boston	8 Denver	13 Memphis	18 Philadelphia	23 Seattle
4 Chicago	9 Detroit	14 Miami	19 Phoenix	24 Tampa
5 Cincinnati	10 Houston	15 Minneapolis	20 Pittsburgh	25 Washington

Fig. 7. The 25 US cities names of the CAB dataset and their approximate locations.



1 Atlanta	**2 Baltimore**	3 Boston	**4 Chicago**	5 Cincinnati
6 Cleveland	**7 Dallas-Fw**	8 Denver	9 Detroit	10 Houston

Fig. 8. Optimal solution to the problem with 10 nodes and three hubs: hub failure consideration. Hub facility.

The optimal/best solutions we report for SApHM-HU and those presented in the literature for the single allocation p-hub median problem [36] show that the topology of the networks provided by the two models differ either with regard to the location of the hubs or the allocation of the non-hubs or both (see for instance, Figs. 1 and 8). Nevertheless, for the high objective weight values (e.g., $w \geq \geq 0.7$) where the minimization of the regular transportation cost has relatively more priority, we expect an increase in similarities between the final solutions provided by the two models namely SApHM and SApHM-HU.

4.5. Practical consideration: the case of multiple backups

In this section we investigate the case where the total demand at a disrupted hub is not necessarily allocated to one single operational hub only. This situation arises especially in the airline industry.

To cater for such a scenario, we adapt our methodology accordingly. We refer to this relaxed problem in which demand points affected by disruption are allowed to be reallocated to any of the operational hubs in the network as the *single hub failure -multiple backups* problem.

4.5.1. An illustrative example of networks with single and multiple backups

Our preliminary analysis shows that the solution topology of a problem with single backup might be different than that of the same problem with multiple backups consideration. To investigate this, we examine the two solutions to a problem with 10 nodes with single and multiple backups considerations.

Fig. 9 compares the topologies of a problem from CAB dataset with 10 nodes, 3 hubs, a discount factor α of 0.2 and a objective weight factor of 0.3 (i.e., 10n3H0.2 α 0.3w) when single and multiple backups are considered. As shown in Fig. 9, the location of hub facilities in both networks is the same but the two networks differ in terms of demand allocation (i.e., node 8 is now assigned to hub 4 instead of hub 7).

With regard to the selected backup hubs, examining the two solutions indicates that when a single backup is considered for the entire demand affected by a disruption, hub 2 is recommended as the backup for hub 4 and hub 4 as the backup for both hub 2 and hub 7. However, when it comes to the case of considering multiple backups, the solution recommends that demand points 1, 5, 6 and 9 to be allocated to hub 2 and hub 7 to act as the backup for node 8. In this case hub 4 is considered as the backup for demand nodes 3 and 10.

In the above example problem, the differences between the corresponding topologies and backups of the two solutions resulted in a cost difference as well. In case of multiple backups, the cost of the best solution found for the problem is 1.07 unit lower than the cost of the optimal solution to the same problem with single backup consideration.

4.5.2. Methodology

The proposed mathematical formulation for the single hub failure-single backup, as described earlier, can be extended to represent multiple backups by implementing some adjustments and introducing new binary variables such as, $u_{i,k,n}$ and $u_{j,k,m}$, to account for backup

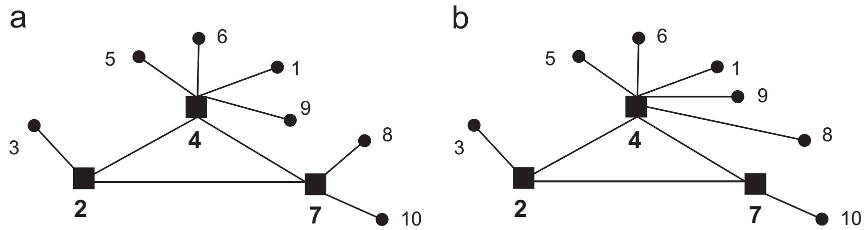


Fig. 9. A network topology comparison between single and multiple backup strategies. (a) Single backup consideration and (b) multiple backups consideration.

Table 6
GA results for the single hub failure multiple backup.

n[a]	p[b]	α[c]	w[d]	Case I:single hub failure-single backup				Case II: single hub failure-multiple backups			
				Hub locations	Backup hubs	TC[e]	Time (s)	Network	Backup hubs	TC[e]	Time (s)
10 (CAB)	3	0.2	0.3	4, 2, 7	2, 4, 4	367.24	0.98	4224447747	2○4○22○424	367.24	3.14
			0.5	4, 2, 7	2, 4, 4	409.89	0.45	4224447747	2○4○22○424	409.89	0.47
			0.7	4, 2, 7	2, 4, 4	452.54	0.64	4224447747	2○4○22○424	452.54	4.96
		0.4	0.3	4, 2, 7	2, 4, 4	401.54	0.72	4224447747	2○4○22○424	401.54	0.87
			0.5	4, 2, 7	2, 4, 4	457.13	0.97	4224447747	2○4○22○424	457.13	6.46
			0.7	4, 2, 7	2, 4, 4	512.72	1.20	4224447747	2○4○22○424	512.72	3.12
		0.8	0.3	4, 2, 7	2, 4, 4	470.15	1.30	4224447447	2○4○22○724	469.08	1.80
			0.5	4, 2, 7	2, 4, 4	551.62	0.62	4224447447	2○4○22○724	550.10	4.60
			0.7	4, 9, 7	9, 4, 4	621.87	1.40	4994497497	944○94○7○9	621.00	1.91
	5	0.2	0.3	1, 2, 4, 7, 8	8, 4, 2, 1, 7	375.46	4.20	1224447847	○○1○22○○21	378.24	1.95
			0.5	1, 2, 4, 7, 8	8, 4, 2, 1, 7	369.16	6.10	1224997797	○○9○42○4○1	373.39	7.59
			0.7	1, 2, 4, 7, 8	8, 4, 2, 1, 4	362.78	3.60	1224447847	○○4○11○○21	364.07	0.85
		0.4	0.3	1, 2, 4, 7, 8	8, 4, 2, 1, 7	432.36	3.20	1224447847	○○4○12○○24	433.69	4.33
			0.5	1, 2, 4, 7, 8	8, 4, 2, 1, 2	437.50	3.00	1224447847	○○4○22○○21	437.50	4.70
			0.7	1, 2, 4, 9, 7	7, 9, 7, 4, 4	443.94	7.60	1224497797	○○9○24○4○4	445.72	7.49
		0.8	0.3	1, 2, 4, 7, 8	7, 4, 2, 1, 2	547.03	2.20	1224447847	○○4○22○○21	547.03	7.9
			0.5	1, 2, 4, 7, 8	2, 4, 2, 1, 2	574.43	5.90	1994497897	○44○94○○○4	575.94	5.88
			0.7	1, 9, 4, 7, 8	8, 4, 8, 1, 7	588.80	2.80	1994497897	○44○94○○○1	589.97	4.35
10 (TR)	3	0.2	0.3	1, 2, 7	2, 1, 1	293.24	5.84	1227127172	○○1121○211	293.24	8.73
			0.5	1, 2, 7	2, 1, 1	316.79	0.30	1227127171	○○1121○212	319.34	2.20
			0.7	1, 2, 7	2, 1, 1	340.35	0.47	1227127171	○○1121○722	344.67	0.71
		0.4	0.3	1, 2, 7	2, 1, 1	331.74	0.50	1227127172	○○1121○211	332.69	3.74
			0.5	1, 2, 7	2, 1, 1	371.96	0.30	1227127171	○○1121○212	373.83	2.66
			0.7	1, 2, 7	2, 1, 1	412.18	0.56	1227127172	○○7121○211	413.96	0.87
		0.8	0.3	1, 2, 7	2, 1, 2	405.72	0.27	1221127111	○○1721○272	404.10	3.53
			0.5	1, 2, 7	2, 1, 1	475.58	0.42	1221127111	○○1721○222	473.68	1.96
			0.7	1, 2, 7	2, 1, 2	543.38	0.61	1221127111	○○1721○272	542.68	3.79
	5	0.2	0.3	1, 2, 4, 7, 8	8, 8, 7, 8, 2	333.59	1.02	1224127848	○○8○88○○72	333.59	5.20
			0.5	1, 2, 4, 7, 8	8, 8, 7, 8, 2	324.55	1.06	1224127848	○○8○88○○82	325.24	6.59
			0.7	1, 2, 4, 6, 7	2, 6, 7, 2, 6	288.63	1.37	1224167146	○○6○2○○211	289.09	2.60
		0.4	0.3	1, 2, 4, 7, 8	8, 8, 1, 8, 2	393.88	0.46	1224127848	○○8○88○○12	394.24	4.78
			0.5	1, 2, 4, 7, 8	8, 8, 1, 8, 2	395.74	3.18	1224127848	○○8○88○○82	396.13	7.05
			0.7	1, 2, 4, 6, 7	2, 6, 1, 1, 6	391.22	2.61	1224167146	○○6○2○○272	391.18	1.55
		0.8	0.3	1, 2, 4, 7, 8	8, 1, 1, 8, 1	513.05	0.90	1224127848	○○1○81○○12	513.77	1.20
			0.5	1, 2, 4, 7, 8	8, 1, 1, 8, 1	537.11	2.50	1224127848	○○1○21○○71	538.93	7.21
			0.7	1, 2, 4, 7, 8	8, 1, 1, 8, 1	561.18	1.64	1224127848	○○8○88○○12	561.53	2.07

[a] Number of nodes.
[b] Number of hubs.
[c] Discount factor.
[d] Objective weight.
[e] Total cost.

hubs for demand points i and j. This modified non-linear and mixed integer problem is significantly more complex than the one developed for the single hub failure-single backup. This problem is larger in terms of both the number of variables and constraints and hence not appropriate for the exact methods as previously shown in the first case namely the single hub failure-single backup.

Our meta-heuristic namely the proposed GA has shown to be powerful and flexible enough to accommodate the necessary changes without incurring a considerable extra computational burden. This flexibility of the GA proved that meta-heuristics are the best way forward in tackling, and adapting to, complex decision problems. The following changes to our initial GA algorithm are made.

In the proposed GA, the fitness value of an offspring for single hub failure-multiple backups case includes three terms. The calculation of the regular transportation cost (i.e., the first term) and the penalty cost of losing the demand at (assumed disrupted) hubs (i.e., the second term) are straight forward; these costs are calculated using the extracted data from the solution representation e.g., locations and allocations data. To calculate the third term (i.e., the rerouting cost), one of the p-hubs in the network is first selected and then a backup facility is assigned (randomly) to each demand point allocated to that particular hub. The entire flow originated from each demand point (initially passes through the disrupted hub) is then rerouted via the backup facilities assigned to that node and the rerouting cost is calculated. The above steps are repeated for all selected hubs in the network and the rerouting cost are summed to represent the third term of the fitness value for the generated offspring. The rest of the steps of the modified GA are similar to that of the initial GA described earlier for the single hub failure-single backup case.

4.5.3. Computational results for the single hub failure-multiple backups case

We tested the new GA on 36 instances with 10 nodes and 3 and 5 hubs. The computational results are presented in Table 6. In this table, the "network" represents the location of the hubs and the allocation of non-hubs to these hubs; "backup hubs" denote the selected backup facilities for all non-hub nodes in the network. For instance, in the first row of Table 6, network 4224447747 shows that nodes 4, 2 and 7 are selected as hubs where nodes 1, 4, 5, 6 and 9 are being allocated to hub 4; nodes 2 and 3 are allocated to hub 2; and nodes 7, 8, and 10 are being assigned to hub 7. The associated "backup hubs" for the described network (i.e., 2∘4∘22∘424) recommend hub 2 as the backup for nodes 1, 5, 6, and 9; and hub 4 as the backup for nodes 3, 8 and 10.

It is worth noting that the optimal solution to the problem with single hub failure multiple backups is either equal or less than that of the single hub failure-single backup case. In other words, the optimal solution to a single hub failure-single backup problem provides an upper bound to the same problem with single hub failure-multiple backups. This is mainly because the solution space of the problem with single hub failure multiple backups is relatively larger and obviously includes the optimal solution of the problem with single hub failure-single backup. Note that this claim is valid only if the problem is solved to optimality.

The above statement is illustrated in our computational results presented in Table 6. Here, in some cases the best solutions to the instances with multiple backups have lower cost compared to that of the same problem with single backup consideration.

Finally, comparisons of the costs and computational times for instances with single and multiple backups in Table 6 also show that the GA has difficulties when solving such problems. This is as one may expect given that the modified GA should deal with a larger solution space and hence a larger number of combinations to consider. According to Table 6, allowing such flexibility provides a cost saving up to 1.9 units.

In summary, our computational results for small problem instances show that (1) the topology of optimal networks for a problem with single and multiple backup considerations might be different in terms location of hubs, demand allocations and/back up selection; (2) in the absence of a constraint to allocate the entire affected demand to a single backup, a network with multiple backups is likely to be less costly than that with single backup consideration; and (3) solving problem instances of the relaxed problem is computationally more expensive than those with single backup consideration.

4.6. GA performance on the classical *p*-hub median problem (special case)

The proposed model for SApHM-HU can be easily tailored for the single allocation p-hub median problem by setting the objective function weight to one (i.e., $w=1$) and removing constraints (15)–(22). A number of problem instances derived from CAB dataset [32] are used as a platform to evaluate the performance of the proposed GA on this special case.

Table 7 summarizes the computational results for the classical single allocation p-hub median problem. Computational time for each problem is fixed at 10 s. The proposed GA solved 11 out of the 24 instances (46%) to optimality in a very short computing time. For the rest of the problems near optimal solutions are also reported. The average computational time to find an optimal/best solution is 2.23 s with an average percentage gap of 0.93.

Table 7
Computational results for single allocation p-hub median problem.

n[a]	p[b]	α[c]	Optimal solution	GA best	Time (s)	%Gap
10	3	0.2	491.93	491.93	0.11	0
		0.4	567.91	567.91	0.16	0
		0.8	716.98	716.98	0.19	0
	4	0.2	395.13	395.13	0.15	0
		0.4	493.79	493.79	0.19	0
		0.8	661.41	661.41	0.16	0
15	3	0.2	799.97	799.97	0.50	0
		0.4	905.10	905.10	0.22	0
		0.8	1099.51	1110.52	0.19	1.0
	4	0.2	639.77	639.77	0.40	0
		0.4	779.71	782.70	0.70	0.4
		0.8	1026.52	1044.45	0.76	1.7
20	3	0.2	724.54	724.54	1.20	0
		0.4	847.77	853.42	3.90	0.7
		0.8	1091.05	1102.34	4.10	1.0
	4	0.2	577.62	585.04	3.98	1.3
		0.4	727.10	727.10	2.19	0
		0.8	1008.49	1036.43	1.40	2.8
25	3	0.2	767.35	786.44	3.86	2.5
		0.4	901.70	909.19	7.88	0.8
		0.8	1158.83	1175.21	8.21	1.4
	4	0.2	629.63	635.82	4.17	1.0
		0.4	787.51	819.67	3.06	4.1
		0.8	1087.66	1126.60	7.57	3.6
Mean			786.96	795.50	2.23	0.93

[a] Number of nodes.
[b] Number of hubs.
[c] Discount factor.

5. Conclusion and future work

In this paper, we present models that incorporate hub unavailability into the classical single allocation p-hub median problem. In this work, we allocate a "backup facility" to each hub centre in the network in order to reduce the effects of hub failures on the operating cost. In the event of hub failure, the demand initially served by the disrupted hub is reallocated to its backup facility. To capture real world situations, we assume that the hub failure probabilities are independent and location specific (i.e., heterogeneous hub failure probabilities). The objective function of the proposed formulation minimizes the weighted sum of regular and expected transportation costs.

We first formulate the problem as a mixed integer quadratic problem followed by our linearization scheme. Though our enhanced formulation is more efficient, it is found to be inappropriate to solve larger instances. To alleviate this drawback, we opted for the design of an evolutionary approach such as GA where a chromosome representation and an efficient crossover operator were specifically developed.

To show the difficulty in solving SApHM-HU using commercial softwares, we solved 36 instances by CPLEX 12.4. Comparison of the results obtained by our GA and those provided by CPLEX shows that the proposed meta-heuristic significantly outperforms CPLEX in computing time while providing the same quality solutions (i.e., optimal solution) for the smaller instances. In addition, 108 larger benchmark problems from CAB [32] and TR [12] datasets are also generated and solved using our efficient GA with promising results. As a special case, we further examined the performance of our algorithm on the conventional single allocation p-hub median problem where our GA-based heuristic provided good quality solutions for the 24 problem instances from the literature while requiring very short computing times.

In brief, we observed that, for moderate and low objective weights, configuration of hubs and spokes in solutions to the single allocation p-hub median problem with and without hub unavailability consideration (i.e., SApHM and SApHM-HU) are different. This finding suggests that a hierarchical solution approach to SApHM-HU (i.e., optimizing the location-allocation problem and then solving the backup assignments problem) may not lead to the same (optimal) solution as the integrated solution approach. Using the problem with 10 nodes, we also show that if hub unavailability is taken into consideration at the design stage, a considerable reduction in expected transportation cost could be achieved without a substantial increase in regular transportation cost. In this study, we also investigated the relaxed problem in which the affected demand nodes are allowed to be reallocated to any of the operational hubs in the network. To this end, the proposed genetic algorithm is adapted to account for multiple backups selection in the case of single hub failure.

Our research paves the path for a thorough study of the impact of hub failure in the context of hub-and-spoke. The problem of multiple backups is an important strategic issue that needs to be considered seriously when locating hubs even though extra inconvenience may arise due to allocating all the disrupted total demand not to one backup hub only. An in depth study for this new logistical problem would be challenging to academics but practically useful to practitioners.

Appendix

City names in TR dataset.

81 nodes problem instance				
1 ADANA	18 ÇANKIRI	35 İZMİR	52 ORDU	69 BAYBURT
2 ADIYAMAN	19 ÇORUM	36 KARS	53 RİZE	70 KARAMAN
3 AFYON	20 DENİZLİ	37 KASTAMONU	54 SAKARYA	71 KIRIKKALE
4 AĞRI	21 DİYARBAKIR	38 KAYSERİ	55 SAMSUN	72 BATMAN
5 AMASYA	22 EDİRNE	39 KIRKLARELİ	56 SİİRT	73 ŞIRNAK
6 ANKARA	23 ELAZIĞ	40 KIRŞEHİR	57 SİNOP	74 BARTIN
7 ANTALYA	24 ERZİNCAN	41 KOCAELİ	58 SİVAS	75 ARDAHAN
8 ARTVİN	25 ERZURUM	42 KONYA	59 TEKİRDAĞ	76 IĞDIR
9 AYDIN	26 ESKİŞEHİR	43 KÜTAHYA	60 TOKAT	77 YALOVA
10 BALIKESİR	27 GAZİANTEP	44 MALATYA	61 TRABZON	78 KARABÜK
11 BİLECİK	28 GİRESUN	45 MANİSA	62 TUNCELİ	79 KİLİS
12 BİNGÖL	29 GÜMÜŞHANE	46 KAHRAMANMARAŞ	63 ŞANLIURFA	80 OSMANİYE
13 BİTLİS	30 HAKKARİ	47 MARDİN	64 UŞAK	81 DÜZCE
14 BOLU	31 HATAY	48 MUĞLA	65 VAN	
15 BURDUR	32 ISPARTA	49 MUŞ	66 YOZGAT	
16 BURSA	33 İÇEL	50 NEVŞEHİR	67 ZONGULDAK	
17 ÇANAKKALE	34 İSTANBUL	51 NİĞDE	68 AKSARAY	

55 nodes problem instance				
1 ADANA	12 BURSA	23 HATAY	34 KAHRAMANMARAŞ	45 AKSARAY
2 ADIYAMAN	13 ÇANAKKALE	24 ISPARTA	35 MARDİN	46 BAYBURT
3 AĞRI	14 ÇORUM	25 İÇEL	36 MUŞ	47 KARAMAN
4 ANKARA	15 DİYARBAKIR	26 İSTANBUL	37 SAMSUN	48 BATMAN
5 ANTALYA	16 EDİRNE	27 İZMİR	38 SİİRT	49 BARTIN
6 ARTVİN	17 ELAZIĞ	28 KAYSERİ	39 SİNOP	50 ARDAHAN
7 BALIKESİR	18 ESKİŞEHİR	29 KIRŞEHİR	40 TEKİRDAĞ	51 IĞDIR
8 BİLECİK	19 GAZİANTEP	30 KOCAELİ	41 TUNCELİ	52 YALOVA
9 BİNGÖL	20 GİRESUN	31 KONYA	42 ŞANLIURFA	53 KARABÜK
10 BİTLİS	21 GÜMÜŞHANE	32 KÜTAHYA	43 YOZGAT	54 KİLİS
11 BURDUR	22 HAKKARİ	33 MANİSA	44 ZONGULDAK	55 OSMANİYE

25 nodes problem instance				
1 ADANA	6 BURSA	11 İÇEL	16 KONYA	21 SAMSUN
2 ADIYAMAN	7 DİYARBAKIR	12 İSTANBUL	17 KÜTAHYA	22 TEKİRDAĞ
3 ANKARA	8 ESKİŞEHİR	13 İZMİR	18 MANİSA	23 ŞANLIURFA
4 ANTALYA	9 GAZİANTEP	14 KAYSERİ	19 KAHRAMANMARAŞ	24 YOZGAT
5 BALIKESİR	10 HATAY	15 KOCAELİ	20 MARDİN	25 ZONGULDAK

10 nodes problem instance	
1 ADANA	6 MANİSA
2 ANTALYA	7 MARDİN
3 BURDUR	8 NEVŞEHİR
4 ELAZIĞ	9 TUNCELİ
5 İÇEL	10 DÜZCE

References

[1] Abdinnour-Helm S. A hybrid heuristic for the uncapacitated hub location problem. Eur J Oper Res 1998;106:489–99.

[2] Alumur S, Kara BY. Network hub location problems: the state of the art. Eur J Oper Resarch 2008;190:1–21.

[3] Alumur S, Kara BY. A hub covering network design problem for cargo applications in Turkey. J Oper Res Soc 2009;60:1349–59.

[4] Berman O, Krass D. On *n*-facility median problem with facilities subject to failure facing uniform demand. Discret Appl Math 2011;159:420–32.

[5] Bryan DL, O'Kelly ME. Hub-and-spoke networks in air transportation: an analytical review. J Reg Sci 1999;39:275–95.

[6] Campbell JF. Integer programming formulations of discrete hub location problems. Eur J Oper Res 1994;72:387–405.

[7] Carello G, Della Croce F, Ghirardi M, Tadei R. Solving the hub location problem in telecommunication network design: a local search approach. Networks 2004;44:94–105.

[8] Cetiner S, Sepil C, Süral H. Hubbing and routing in postal delivery systems. Ann Oper Res 2010;181:109–24.

[9] Church RL, Scaparra MP, Middleton RS. Identifying critical infrastructure: the median and covering facility interdiction problems. Ann Assoc Am Geogr 2004;94:491–502.

[10] Cui T, Ouyang Y, Shen Z-JM. Reliable facility location design under the risk of disruptions. Oper Res 2010;58:998–1011.

[11] Cunha CB, Silva MR. A genetic algorithm for the problem of configuring a hub-and-spoke network for a LTL trucking company in Brazil. Europ J Oper Res 2007;179:747–58.

[12] Çetiner S. An iterative hub location and routing problem for postal delivery systems [Mater's thesis]. Department of Industrial Engineering, Middle East Technical University, Ankara, Turkey; 2003.

[13] Ebery J. Solving large single allocation *p*-hub problems with two or three hubs. Eur J Oper Res 2001;128:447–58.

[14] Eiselt HA, Gendreau M, Laporte G. Optimal location of facilities on a network with an unreliable node or link. Inf Process Lett 1996;58:71–4.

[15] Ernst AT, Krishnamoorthy M. Efficient algorithms for the uncapacitated single allocation *p*-hub median problem. Locat Sci 1996;4:139–54.

[16] Glover F. Future paths for integer programming and links to artificial intelligence. Comput Oper Res 1986;13:533–49.

[17] Holland J. Adaptation in natural and artificial systems. Ann Arbor: University of Michigan Press; 1975.

[18] Janic M. Modelling the large scale disruption of an airline network. J Transp Eng 2005;131:249–60.

[19] Jeong SJ, Lee CG, Bookbinder JH. The European freight railway system as a hub-and-spoke network. Transp Res Part A 2007;41:523–36.

[20] Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. Science 1983;220:671–80.

[21] Kim H, O'Kelly ME. Reliable p-hub location problems in telecommunication networks. Geogr Anal 2009;41:283–306.

[22] Klincewicz JG. Hub location in backbone/tributary network design: a review. Locat Sci 1998;6:307–35.

[23] Kratica J, Stanimirović Z, Tošić D, Filipović V. Two genetic algorithms for solving the uncapacitated single allocation *p*-hub median problem. Eur J Oper Res 2007;182:15–28.

[24] Kuby MJ, Gray RG. The hub network design problem with stopovers and feeders: the case of Federal Express. Transp Res A 1993;27:1–12.

[25] Lapierre SD, Ruiz AB, Soriano P. Designing distribution networks: formulations and solution heuristic. Transp Sci 2004;38:174–87.

[26] Li Q, Shen M, Snyder LV. The effect of supply disruptions on supply chain design decisions. Transp Sci 2010;44:274–89.

[27] Li Q, Zeng B, Savachkin A. Reliable facility location design under disruptions. Comput Oper Res 2013;40:901–9.

[28] Marianov V, Serra D. Location models for airline hubs behaving as M/D/c queues. Comput Oper Res 2003;30:983–1003.

[29] Martin JC, Roman C. Analysing competition for hub location in intercontinental aviation markets. Transp Res Part E 2004;40:135–50.

[30] Nagy G, Salhi S. Location-routing: issues, models and methods. Eur J Oper Res 2007;177:649–72.

[31] Nickel S, Schobel A, Sonneborn T. Hub location problems in urban traffic networks. In: Niittymaki J, Pursula M, editors. Mathematics methods and optimization in transportation systems. Kluwer Academic Publishers; 2001 (Chapter 1).

[32] O'Kelly ME. A quadratic integer program for the location of interacting hub facilities. Eur J Oper Res 1987;32:393–404.

[33] O'Kelly ME. The location of interacting hub facilities. Transp Sci 1986;20:92–105.

[34] Peng P, Snyder LV, Liu Z, Lim A. Design of reliable logistics networks with facility disruptions. Transp Res Part B: Methodol 2011;45:1190–211.

[35] Sasaki M, Suzuki A, Drezner Z. On the selection of hub airports for an airline hub-and-spoke system. Comput Oper Res 1999;26:1411–22.

[36] Skorin-Kapov D, Skorin-Kapov J, O'Kelly M. Tight linear programming relaxations of uncapacitated p-hub median problems. Eur J Oper Res 1996;94:582–93.

[37] Snyder LV, Daskin MS. Reliability models for facility location: the expected failure cost case. Transp Sci 2005;39:400–16.

[38] Topcuoglu H, Coruta F, Ermisb M, Yilmaza G. Solving the uncapacitated hub location problem using genetic algorithms. Comput Oper Res 2005;32:967–84.

[39] Yaman H, Kara BY, Tansel BC. The latest arrival hub location problem for cargo delivery systems with stopovers. Transp Res B 2007;41:906–19.