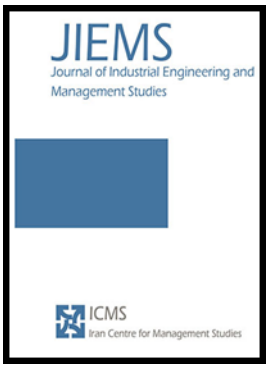
JIEMS
Journal of Industrial Engineering and Management Studies

Vol. 7, No. 2, 2020, pp. 139-165

DOI: 10.22116/JIEMS.2020.195853.1293

www.jiems.icms.ac.ir



Using revenue management technique to allocate the capacity in reliable hub network design under uncertain air passenger traffic

Mojtaba Salehi[1,*], Hamid Tikani[2]

Abstract
This paper introduces a two stage stochastic programming to address strategic hub location decisions and tactical flight routes decisions for various customer classes considering uncertainty in demands. We considered the airline network with the arc capacitated single hub location problem based on complete–star p-hub network. In fact, the flight routes are allowed to stop at most two different hubs. The first stage of the model (strategic level) determines the network configuration, which does not change in a short space of time. The second stage is dedicated to specify a service network consists of determining the flight routes and providing booking limits for all itineraries and fare classes after realization of uncertain scenarios. To deal with the demands uncertainty, a stochastic variations caused by seasonally passengers' demands through a number of scenarios is considered. Since airline transportation networks may face different disruptions in both airport hubs and communication links (for example due to the severe weather), proposed model controls the minimum reliability for the network structure. Due to the computational complexity of the resulted model, a hybrid algorithm improved by a caching technique based on genetic operators is provided to find a near optimal solution for the problem. Numerical experiments are carried out on the Turkish network data set. The performance of the solutions obtained by the proposed algorithm is compared with the pure GA and Particle Swarm Optimization (PSO) in terms of the computational time requirements and solution quality.

Keywords: customer segmentation; scenario generation method; network reliability; stochastic programming; meta-heuristic algorithms.

Received: July 2019-26

Revised: May 2020-30

Accepted: August 2020-31

1. Introduction

Revenue management has emerged on the capacity management by booking classes for available seats in the airline industry. RM techniques (such as seat inventory control, pricing, forecasting and etc.) are arose out after the deregulation in the airline industry and have been considered as powerful tools for maximizing the total revenue (M.G. Yoon et al., 2017). Deregulation in the airline industry makes it is possible for airlines to sell the seats of an

* Corresponding author; m_salehi61@yahoo.com
[1] Payame Noor University, Tehran, Iran.
[2] K. N. Toosi University of Technology, Tehran, Iran.

aircraft to different customer segments at different prices during the booking period. Airlines define some conditions and restrictions as fences to limit switching of the customers among various segments of the market and offer multiple services with various prices. In particular, airline companies try to sell high-fare ticket prices as many seats as possible and attempt to minimize the potential loss that happens as an effect of unsold seats. In this process, they may reject an early customer with a lower-fare ticket so as to save the seat for the customer with higher-fare ticket prices, but at the same time they confront with the risk of flying with some empty seats. Under this circumstance, the airline company has fixed and perishable set of assets that should be sold to a price-sensitive population of customers with the purpose of maximizing the total expected revenue over the selling horizon (Aslani et al. 2014).

Most airlines utilize hub-and-spoke type of network to serve more origin–destination (OD) pairs with fewer flights. Hub and spoke networks cause to receive much interest as the fight legs are shared by multiple origin-destination itineraries. Most of the time, using hub and spoke networks results in lower cost carriers as compared to those of direct flights. Applying optimal inventory control strategies together with choosing the best network topology help the airlines to improve the total expected revenue. This topology uses switching, transshipment and flow consolidation facilities named hubs that remarkably decrease the links required to connect all origins and destinations in transportation networks (Alumur et al., 2012). In other words, consolidating traffic flow in inter-hub transportation and on the spokes causes to reduce the operational costs in comparison with direct links between all pairs and leads to maximize the total revenue.

In the current paper, we extend a model that incorporates uncertainty into the hub location problem in which we allocate the fixed amount of capacity to the right customers with the purpose of maximizing the total revenue that incurred by transfer activities in a reliable hub network design. We consider a hub location problem that arises in the design of a complete–star network. In a complete–star p-hub network, there are several nodes in the network that p of them are chosen to be hubs. Each node is assigned to exactly one hub and all of the hubs are connected to each other. To capture the uncertainty of demand, a two-stage stochastic integer-programming model is formulated. The first stage aims to find the optimal locations of p hubs, the allocation of non-hub nodes to the p located hubs considering the reliability of network, and the second stage provides the decisions on the sold tickets, which are influenced by uncertain demands. Due to the NP-hardness of the problem, to solve the proposed model, we are proposing an algorithm based upon the evolutionary genetic algorithm and exact solution method. The rest of the paper is organized as follows:

Section 2 discusses the relevant literature briefly. The model is developed in Section 3. In Section 4, we provided a hybrid algorithm improved by a caching technique based on genetic operators to solve the problem. A computational study with different numbers of hubs and nodes using Turkish network data set is presented in Section 5. We also analysis the performance of proposed algorithm in section 5.4. Moreover, two well-known ratios for stochastic optimization, the Expected Value of Perfect Information (EVPI) and the Value of the Stochastic Solution (VSS) are calculated to evaluate the developed model section 5.5. Finally, in Section 6, general conclusions and some suggestions for future research are given.

2. Literature review

2.1. Revenue management problem

Revenue Management (RM) has attracted the attention of practitioners and researchers for many years. The methods and techniques in this science attempt to maximize the revenue. Forecasting, pricing and seat controls have played key roles in airlines among different RM techniques (M.G. Yoon et al. 2017). Littlewood (1972) is known as a pioneer of revenue

management. He introduced a model to manage the capacity of single flight leg for two fare classes.

There is a considerable amount of literature regarding revenue management in the airline industry. Belobaba (1989) applied Littlewood's (1972) rule and introduced effective heuristic known as EMSR (expected marginal seat revenue), for single leg problems with multi independent demand for pairs of fare classes. He also provided a modified EMSR method, called EMSR-b (Belobaba and Weatherford (1996)). EMSR-b differs in aggregating the demand of the products for reducing the problem with p fare classes to two-class problems based on a weighted combination in which seat protection levels near to optimal values better than previous EMSR. The problem of seat allocation in the form of multiple fare classes has been studied by many researchers. Brumelle and McGill (1993) provided several static models with different distribution of the demand that determine the optimal booking limits for different fare classes in a single-leg problem. A two-class dynamic seat allocation model with passenger diversion is considered by Zhao and Zheng (2001). Their studies have two particular characteristics, in proposed model discount fare cannot be reopened when it closed, the other feature is that they considered some flexible customers, these customers would buy discount fare tickets if available but willing to pay the full fare. Chen et al. (2010) analyzed the optimal policy for the two-flight and optimal booking problem. Nechval et al. (2013) considered the allocation of the finite seat inventory to the uncertain airline customer demand that occurs during the time before the flight is scheduled to depart. The purpose of proposed model is to maximize the revenue by finding the right combination of customers of various fare classes. Cizaire and Belobaba (2013) proposed a joint approach to optimize fares and booking limits so as to maximize the total revenues generated by the two fare products over the two time periods. Mou and Wang (2014) presented an uncertain programming for network revenue management in which the fares and the demands are considered as uncertain variables.

We refer to study Çetiner (2013), Lapp and Weatherford (2014), and Brumelle and McGill (1993) for an overview on the seat allocation problem and the early researches on this subject.

In trying to offer a more comprehensive framework for revenue management, the performance of a given system, significantly depends on designing optimal networks for routing the traffic. To provide organized transportation between different origins and destinations, utilizing set of hubs and using fewer arcs instead of point-to-point network can reduce transportation costs.

2.2. Reliable hub location problem

In HLP networks, any failure or malfunction in components may cause inconstancy and degradation of the entire network's capability to transfer flows (Kim and O'Kelly 2009). Kim and O'Kelly (2009) utilized the concept of reliability in the HLP for the first time and introduced a reliable p-HLP which focused on maximizing network performance in terms of reliability by locating hubs for delivering flows among origin destination pairs.

Davari et al. (2010) addressed a single-allocation hub-and-spoke network with the aim of maximizing the reliability of the network. They assumed that the reliability of each arc is a fuzzy variable. Fazel Zarandi et al. (2011) presented a reliable single-allocation hub-and-spoke network using an interactive fuzzy goal programming. Azizi et al. (2016) presented a mathematical model that builds hub-and-spoke systems under the risk of hub disruption. They assume that once a hub stops normal operations, the entire demand initially served by this hub is handled by a backup facility. An et al. (2015) introduced a set of reliable hub and-spoke network design models, where the selection of backup hubs and alternative routes are taken into consideration to handle hub disruptions. Eghbali et al. (2014) studied multi-

objective single hub location problem, where all the routes that utilized for transmitting the traffic flow are controlled to have minimum reliability.
Numerous studies have attempted to solve the p-hub location problem with genetic algorithm. Topcuoglu et al (2005) presented a robust solution based on a genetic algorithm search for the uncapacitated single allocation hub location problem called USAHLP, and numerical experiments were carried out on the CAB and AP data set for solving the problem. Kratica et al (2007) proposed two genetic algorithm (GA) approaches improved by caching strategy for solving an uncapacitated single allocation p-hub median problem. Damgacioglu et al (2014) proposed a genetic algorithm approach to solve a single hub location problem with an uncapacitated and planar model called PHLP in a reasonable amount of time. Momayezi et al. (2018) developed a capacitated modular with single assignment considered backup for failed hubs.

2.3. Uncertainty of demands in hub location problem

In a real-world application in the airline market, demand has some uncertainty during the time. The importance of uncertainty has motivated several researchers to study various stochastic parameters in network design problems. Sim et al. (2009) considered chance constraints in the stochastic p-hub center problem to formulate the service-level guarantees. The only source of uncertainty in their study is travel time. Alumur et al. (2012) proposed the hub location problem by considering two sources of uncertainty contains the set-up costs for the hubs and the demands between origin–destination pairs. Contreras et al. (2011) addressed stochastic uncapacitated hub location problems considering uncertainty in demands and transportation costs. Yang (2009) and Yang and Chiu (2016) presented a stochastic programming model to determine the air freight hub location and flight routes planning in the stochastic environment. Adibi and Razmi (2015) developed an uncapacitated multiple hub location problem with stochastic demand and transportation cost. They used 10-node air network in Iran to evaluate their proposed model.

2.4. An integrated hub location and revenue management problem

Hub location is one of the most attractive fields in facility location which has been appeared in various applications in the real world, including airline systems, postal delivery systems, cargo delivery systems and telecommunication network design. Since the service network design plays an important role in airline operations, many researchers have been addressed this problem. O’Kelly (1987) was the first one who formulated the discrete hub location problem as a quadratic integer program. After that, this field have attracted attentions of researchers. The large, and growing researches, solution techniques and applications on hub location problem is summarized in (Alumur and Kara 2008), (Farahani et al. 2013) and (Campbell and O'Kelly 2012).
The purpose of the classical hub location is to minimize the total cost including establishment cost and the cost of traffic flows under the assumption of serving all demands. However, in order to maximize the profit, we can serve only a portion of flows between each origin-destination pair. Accordingly, as a relatively novel approach, some studies in the literature discussed the trade-off between the obtained revenue and the total costs of the transportation system. We can classify the researches that incorporate the profit maximization into hub location problem as follows:

- Some of the studies including Alibeyg et al. (2016), Neamatian Monemi et al. (2017), Alibeyg et al. (2018), Taherkhani and Alumur (2019) only focused on network design problem to captured profitable flows between demand nodes.

- While other papers like Lüer-Villagra and Marianov (2013), Lin and Lee (2018), and Čvokić and Stanimirović (2020) incorporated pricing decisions into hub location problem.
- Finally, some of the studies like Tikani et al. (2018) and Huo et al. (2019a) not only considered network design to captured profitable flows but also focused on integration of RM techniques and hub location problem (RM-HLP).

The current paper is dedicated to the latter and addresses an integrated model of RM and HLP. Firstly, this problem is introduced by Tikani et al. (2016). They studied the seat inventory control decisions in a hybrid transportation network considering both hub-stop flights and non-stop flights. Tikani et al. (2018) formulated the problem of Tikani et al. (2016) in a complete in a star/star hub network. In the following, Huo et al. (2019a) analyzed the RM-HLP with a weighted sum function including average and worst-case profits. Then, Huo et al. (2019b) under multiple capacity levels for hubs. To the best of author's knowledge from a review of the literature, this is the first work that studies the reliability of hubs and links in RM-HLP. The problem is named as the reliable integrated hub location and revenue management problem abbreviated by R-RM-HLP. In detail, R-RM-HLP considers the entire network's ability based on Kim and O'Kelly (2009) and Eghbali et al. (2016) to increase customer convenience in an airline RM system. It also focuses with seat allocation decisions in flight routes based upon stochastic programming and the scenario generation method. In the current study, we provided a new formulation in order to maximizes airline's profit by designing best capacitated reliable topology network and routing policies by applying optimal hub and spoke, then optimal seat allocation for different customer classes is carry out for each itinerary regarding to available capacity. We also propose an efficient hybrid algorithms for solving the problem. In the Table 1, the innovation of the current study is compared to the existing literature.

3. Problem description of R-RM-HLP

The aim of the presented model is to find the location of hubs, the optimal flight routes and the number of seats to allocate on each rout for different classes to maximize the revenue. We consider a hub location problem that arises in the design of a complete–star network. There are several nodes in the network that p of them are chosen to be hubs. Each node is assigned to exactly one hub and all of the hubs are connected to each other. Our hub and spoke network configuration is an extended case of flow-based models which proposed by Adibi and Razmi (2015). We formulate capacitated version of the problem and incorporate customer segmentation strategy in the second stage of the model.

In Figure 1, a complete–star network is depicted. Consider that we have to transfer passengers between two nodes 8 and 9, if both nodes are assigned to the same hub like 2, the passengers from node 8 to node 9 first goes to hub 2 and then from hub 2 goes to node 9 (one-hub-stop flights). In a situation that node 7 and node 13 are assigned to different hubs the passengers traffic between these nodes first goes from node 7 to its hub like 3, and then from hub 3 to the hub 1 and lastly the traffic goes from hub 1 to node 13. The limited capacity on arc (i, j) is shared between itineraries from i to the other destination nodes and the itineraries between all origin nodes to the node i. Therefore, we have two types of flight legs in the network. The first one is the links that transfers the traffic between non hub node and a hub node that we name them as link type 1 and the second type is that transfers the traffic between hubs that we name them as links type 2. The classical discount factor α is incorporated to the model by taking $C2 = \alpha\, C1$. It is assumed customers are classified into different segments or classes based on their sensitivity to prices. Considering the locations of

hubs and the fixed and limited capacity for serving uncertain demand, the sale decisions are made after the arrival of customer orders as well as when more definitive information becomes available.

The decisions about network structures in hub location problems such as locating hubs and the allocation of non-hub nodes to the hubs is substantially important due to the major capital investments to establish them and since the transportation system must be utilized for many years in the future (Eghbali et al. 2014). The operation of each element (hub and arc) in the system affect the overall performance of the network acutely. In air transportations, airlines have to face various disruptions such as severe weather, labor strikes, terrorism threats, and runway incursions disrupt regular operations that make airports partially or completely unavailable (An et al. 2015). Such problems may cause delaying, canceling and etc. that bring customer dissatisfaction.
The general assumptions of the hub location problem (which are in common with R-RM-HLP) are:

- The number of hubs is predefined.
- The location of hub nodes should be specified by the model.
- Direct connection between two non-hub nodes is prevented.
- Each non-hub is only allowed to be assigned to exactly one hub and all of the hubs are connected by a complete network
- The traffic flows between two hub nodes (by leg type 2) are discounted by parameter due to the economies of scale.

Table 1. Related works on profit maximization for hub location problem and their solution methods

Authors	The strategy of profit maximization			Uncertainty	Reliability aspects		Hub network structure	Application/Data set	Solution approach
	network design	pricing decisions	RM techniques		Reliability of arcs	Reliability of hubs			
Kim and O'Kelly (2009)	✗	✗	✗	✗	✓	✓	Complete-star	CAB data set	CPLEX solver
Lüer-Villagra and Marianov, (2013)	✗	✓	✗	✗	✗	✗	Complete-star	CAB data set	Genetic algorithm
Eghbali et al. (2014)	✗	✗	✗	✗	✓	✓	Complete-star	Turkish dataset	non-dominated sorting genetic algorithm-II
Tikani et al. (2016)	✓	✗	✓	✓	✗	✗	Star-star	Turkish dataset	Genetic algorithm
Alibeyg et al. (2016)	✓	✗	✗	✗	✗	✗	incomplete	CAB data	CPLEX solver
Neamatian Monemi et al. (2017)	✓	✗	✗	✗	✗	✗	Complete-star	real data	Metaheuristic approach combining a local search algorithm and a Lagrangian relaxation
Lin and Lee (2018)	✗	✓	✗	✗	✗	✗	Complete-star	Cities in Taiwan	Lagrangian relaxation
Tikani et al. (2018)	✓	✗	✓	✓	✗	✗	Star-star	Turkish dataset	Genetic algorithm
Alibeyg et al. (2018)	✓	✗	✗	✗	✗	✗	incomplete	CAB data	Lagrangian relaxation
Huo et al. (2019a)	✓	✗	✓	✓	✗	✗	Star-star	Turkish dataset	Genetic algorithm
Taherkhani and Alumur (2019)	✓	✗	✗	✗	✗	✗	Complete-star	CAB data	CPLEX solver
Huo et al. (2019b)	✓	✗	✓	✓	✗	✗	Star-star	Generated instances	Genetic algorithm
Čvokić and Stanimirović (2020)	✗	✓	✗	✓	✗	✗	Complete-star	CAB dataset	CPLEX solver
Current study	✓	✗	✓	✓	✓	✓	Complete-star	Turkish dataset	Particle swarm optimization, Modified Genetic algorithm, Standard GA, CPLEX solver

Moreover, the specific assumptions of R-RM-HLP are listed below.

- The employed aircrafts through leg type 1 and type 2 have limited by capacity Q1 and Q2, respectively
- The price of the ticket for each fare class is known in advance.
- Cancellations are not permitted. Moreover, upgrading or downgrading the ticket class or changing the departure time of flight is not allowed.
- Upgrading to a higher fare class or changing to a later or earlier flight shouldn't be allowed.
- Traffic demands for each type of fare class k is considered to be independent.
- The air traffic passenger for each class is uncertain. The booking limit is related to both the protection level of tickets and the uncertain traffic passenger.
- It is not necessary to satisfy all the demands between origin-destination pairs and the objective function is calculated based on the number of tickets sold on the profitable itineraries.

As mentioned before, network design is determined at the beginning of the horizon. In addition, the sale decisions are made after the arrival of customer orders as well as when more definitive information becomes available. On the other hand, the customer classes arrive with stochastic demands. The tickets are sold according to both the capacities and demands to maximize the total airline's revenue. In this section, first we introduce the non-linear revenue maximization and capacitated hub location model with different customer classes. In the second part of this section, a linear Deterministic Equivalent Problem is obtained by some additional variables and constraints.
Notations used for mathematical formulation are as follows:

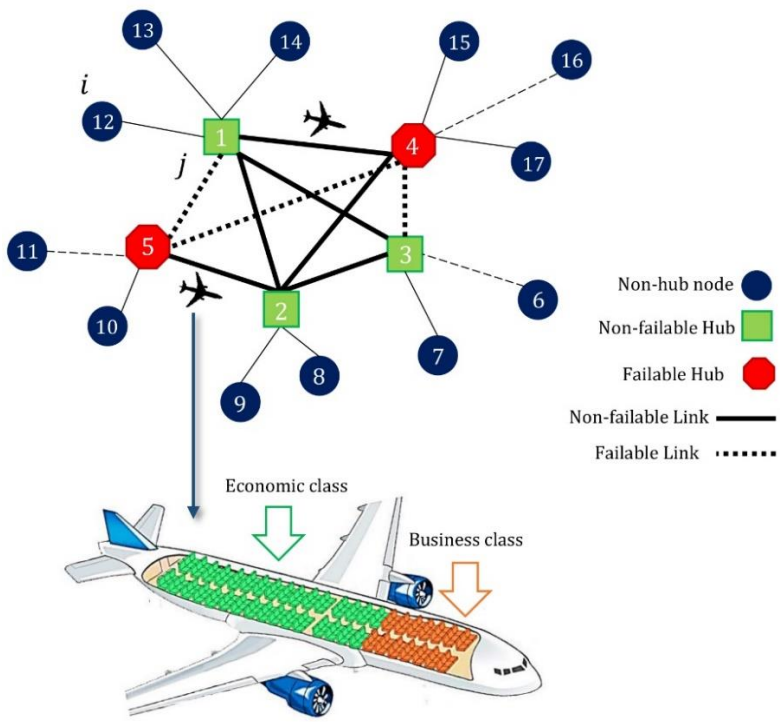


Figure 1. An example solution of presented hub location and seat allocation problem

3.1. Sets and Parameters

N	The number of nodes (spokes) in the hub-spoke network.
P	The number of hubs.
Ω	The number of scenarios.
K	The number of customer classes.
i, m	Indices for nodes $i, m = 1 \dots N$.
j, l	Indices for hubs.
ω	Indices for scenarios.
k	Indices for customer classes.
f_{im}	Distance from node i to node m.
C_{1k}	Transfer cost (per unit flight and unit distance) between origin and hub node which is defined as leg type 1 for customer class k.
C_{2k}	Transfer cost (per unit flight and unit distance) between a two hubs which is defined as leg type 2 for customer class k.
$D_{im}^{k}(\omega)$	Traffic demand between origin i and destination m for customer class k under scenario ω.
$p(\omega)$	The probability of traffic demand between origin i and destination m for customer class k under scenario ω
r_{im}^{k}	Ticket price of itinerary between origin i and destination m for customer class k.
No_{ij}^{1}	Number of flights available for itinerary between node i and hub j.
No_{jl}^{2}	Number of flights available for itinerary between hub j and hub l.
S_1	Number of seats available in a service cabin at leg type 1.
S_2	Number of seats available in a service cabin at leg type 2.
F_j	Fixed cost for establishing hub j.
Re_{ij}	Reliability of link between node i and node j
re_j	Reliability of hub j
H	Minimum reliability that considered for each origin/ destination route

3.2. Decision variables

$y_{im}^{k}(\omega)$	Number of tickets sold for itinerary between origin i and destination m for customer class k under scenario ω.
x_{ijlm}	A binary decision variable, which is 1 if traffic from node i to node m goes through hubs located at node j and l and 0 otherwise.
w_{ij}	A binary decision variable, which is 1 if node i is connected to hub j and 0 otherwise.
w_{jj}	A binary decision variable, which is 1 if node j is selected to be hub and 0 otherwise.

With these set of variables, we can obtain a nonlinear formulation as follows

3.3. Non-linear formulation of R-RM-HLP

The nonlinear formulation of the proposed hub location and revenue maximization problem is as follows:

Stage (1):

$$\text{Max } E\left[Q\left(w, \xi(\omega)\right)\right] - \sum_j F_j w_{jj} \tag{1}$$

Subject to

$$\sum_j \sum_l x_{ijlm} = 1 \quad \forall i, m \tag{2}$$

$$x_{ijlm} \leq w_{ij} \quad \forall i, j, l, m \tag{3}$$

$$x_{ijlm} \leq w_{ml} \quad \forall i,j,l,m \tag{4}$$
$$w_{ij} \leq w_{jj} \quad \forall i,j \tag{5}$$
$$\sum_j w_{ij} = 1 \quad \forall i \tag{6}$$
$$\sum_j w_{jj} = P \tag{7}$$
$$Re_{ij} re_j^{1-\beta} Re_{jl}^{1-\gamma} re_l^{1-\beta} Re_{lm} \geq H.x_{ijlm} \quad \forall i,j,l,m \tag{8}$$
$$w_{ij} \in \{0,1\} \quad \forall i,j \tag{9}$$
$$x_{ijlm} \in \{0,1\} \quad \forall i,j,l,m \tag{10}$$

Where $Q\big(w,\xi(\omega_D)\big)$ is the optimal value of the following second stage problem:
Stage (2):

$$Q\big(w,\xi(\omega)\big) = Max \ \sum_i \sum_m \sum_k y_{imk}(\omega) r_{im}^k - \sum_i \sum_j \sum_k C_{1k}[\, f_{ij}(\sum_{m\in I/\{i\}} \frac{y_{im}^k(\omega)}{S_1}) + f_{ji}(\sum_{m\in I/\{i\}} \frac{y_{mi}^k(\omega)}{S_1}))]w_{ij} - \sum_k \sum_j \sum_l \sum_i \sum_{m\in I/\{i\}} C_{2k}\left[f_{jl}\left(\frac{y_{im}^k(\omega)}{S_2}\right) + f_{lj}\left(\frac{y_{mi}^k(\omega)}{S_2}\right)\right] x_{ijlm} \tag{11}$$
$$y_{im}^k(\omega) \leq D_{im}^k(\omega) \quad \forall i,m,k \tag{12}$$
$$\sum_i \sum_m \sum_k \frac{y_{im}^k(\omega)}{S_2} x_{ijlm} + \frac{y_{im}^k(\omega)}{S_2} x_{mjli} \leq NO_{jl}^1 \quad \forall j,l \tag{13}$$
$$\sum_m \sum_k \frac{y_{im}^k(\omega)}{S_1} + \sum_m \sum_k \frac{y_{im}^k(\omega)}{S_1} \leq \sum_j NO_{ij}^2 . w_{ij} + M\, w_{ii} \quad \forall i \tag{14}$$
$$y_{imk}(\omega) \quad \forall i,m,k \tag{15}$$

Problems stage (1) and (2) combine together to achieve a two-stage stochastic program.
The objective function (1) maximizes the total revenue of the airline transportation system and contains a deterministic term$\sum_j F_j w_{jj}$, which is related to fix hub setup costs, and the expectation of the second stage objective $Q\big(w,\xi(\omega_D)\big)$ which is obtained after realizations of random event ω and determine the total revenue (objective function 12). In the other words, the strategic decisions is made in the first stage of the problem. These decisions are relevant to locating hubs and the allocation of non-hub nodes to hubs, which need major capital investments since the transportation network is a long-term plan and should be employed for many years. In fact, network configuration is prior to realizations of the random parameters $\xi(\omega)$.
The second stage of the problem establish the network through the determination of flight routes and booking limits for all itineraries and fare classes (i.e. $y_{im}^k(\omega)$) after the after realization of uncertain scenarios and the first stage decision,w_{ij} has been taken. To put it another way, once the value of w_{ij} and consequently x_{ijlm} is taken value in the first stage, the second stage decisions are may change under various realizations of ω.
Constraint (2) ensures that each origin/destination pair (i, m) is allocated to one pair of hub nodes (j, l). Constraints (3) and (4) guarantee that the demand from origin node i to destination node m cannot be allocated to a hub pair (j, l) unless both nodes (j, l) are defined as hub nodes. Constraint (5) assures that a node can be assigned to a location if a hub is opened at that site. Constraint (6) enforces single allocation for each node. Constraint (7) states that there must be exactly P hubs. Constraint (8) controls the reliability in such way that from origin i to destination m via hubs k and l must not be less than a specified value. Constraint s (9) and (10) guarantee the variables are binary.
The objective function (11) maximizes the total revenue of selling tickets to different classes and itineraries considering total transportation cost between all non-hub nodes and the hub nodes, total transportation cost between hubs.

Constraint (12) insures that the number of sold tickets for each origin destination itineraries should be less than the demands. Constraint (13) establish the capacity constraints of link type 1 between hub node j and hub node l. Also, constraint (14) states the capacity limit of link type 2 between node i and hub node j. Constraint (15) defines the types of decision variables.

3.4. Linear Deterministic Equivalent Problem

The two-stage stochastic programming is computationally hard to be solved by general optimization tools. To employ the stochastic programming (SP) model the demand $D(\omega)$ is assumed to follow a discrete scenarios under seasonal demand variation and the problem can be formulated with a finite number of possible scenarios. So in this section we rewritten the Deterministic Equivalent Problem (DEP) for 2-stage stochastic programming. Proposed DEP of model is a mixed integer problem and equivalent to the original version of the model and on the contrary much simpler to solve. In order to obtain the linear integer programming we add some variables and a set of linear constraints. Now with this explanation linear DEP formulation of the problem is as follows:

$$Max\ \sum_{i\epsilon N}\sum_{m/\{i\}\epsilon N}\sum_{k\epsilon K}\sum_{\omega\epsilon\Omega} p(\omega)y_{im}^{k}(\omega)r_{imk} - \sum_{i\epsilon N}\sum_{m/\{i\}\epsilon N}\sum_{k\epsilon K}\sum_{\omega\epsilon\Omega} C_{1k}[f_{ij}\left(p(w)(\frac{V_{ijm}^{k}(\omega)}{S_1}+\frac{V_{jim}^{k}(\omega)}{S_1})\right)]- \sum_{i\epsilon N}\sum_{m/\{i\}\epsilon N}\sum_{k\epsilon K}\sum_{j\epsilon N}\sum_{l\epsilon N} C_{2k}\sum_{\omega\epsilon\Omega}[f_{jl}\left(p(w)(\frac{O_{ijlm}^{k}(\omega)}{S_1}+\frac{O_{mlji}^{k}(\omega)}{S_1})\right)] - \sum_j F_j w_{jj} \quad (16)$$

$$\sum_j\sum_l x_{ijlm} = 1 \quad \forall i,m \quad (2)$$

$$x_{ijlm} \le w_{ij} \quad \forall i,j,l,m \quad (3)$$

$$x_{ijlm} \le w_{ml} \quad \forall i,j,l,m \quad (4)$$

$$w_{ij} \le w_{jj} \quad \forall i,j \quad (5)$$

$$\sum_j w_{ij} = 1 \quad \forall i \quad (6)$$

$$\sum_j w_{jj} = P \quad (7)$$

$$y_{im}^{k}(\omega) \le D_{imk}(\omega) \quad \forall i,m,k,\omega \quad (12)$$

$$\sum_i\sum_m\sum_k \frac{O_{ijlm}^{k}(\omega)}{S_2} + \sum_i\sum_m\sum_k \frac{O_{mjli}^{k}(\omega)}{S_2} \le NO_{jl}^{1} \quad \forall j,l,\omega \quad (17)$$

$$\sum_m\sum_k \frac{y_{im}^{k}(\omega)}{S_1} + \sum_m\sum_k \frac{y_{mi}^{k}(\omega)}{S_1} \le \sum_j NO_{ij}^{2}.w_{ij} + M\, w_{ii} \quad \forall i,\omega \quad (15)$$

$$Re_{ij}re_j^{1-\beta}Re_{jl}^{1-\gamma}re_l^{1-\beta}Re_{lm} \ge H.x_{ijlm} \quad \forall i,j,l,m \quad (8)$$

$$V_{ijm}^{k}(\omega) \ge (w_{ij}-1)M + y_{im}^{k}(\omega) \quad \forall i,m,j,k,\omega \quad (18)$$

$$V_{mji}^{k}(\omega) \ge (w_{ij}-1)M + y_{mi}^{k}(\omega) \quad \forall i,m,j,k,\omega \quad (19)$$

$$O_{ijlm}^{k}(\omega) \ge (x_{ijlm}-1)M + y_{im}^{k}(\omega) \quad \forall i,m,j,l,k,\omega \quad (20)$$

$$O_{mjli}^{k}(\omega) \ge (x_{mjli}-1)M + y_{im}^{k}(\omega) \quad \forall i,m,j,l,k,\omega \quad (21)$$

$$O_{ijlm}^{k}(\omega), V_{ijm}^{k}(\omega) \ge 0 \quad \forall i,j,l,m,k,\omega \quad (22)$$

$$w_{ij} \in \{0,1\} \quad \forall i,j \quad (9)$$

$$x_{ijlm} \in \{0,1\} \quad \forall i,j,l,m \quad (10)$$

$$y_{im}^{k}(\omega) \in Z_+ \quad \forall i,m,k,\omega \quad (17)$$

By employing constraint (18), the variable $y_{im}^{k}(\omega)$becomes the lower bound for $V_{ijm}^{k}(\omega)$, if w_{ij} takes value of 1, and using constraint (20), the variable $y_{im}^{k}(\omega)$ becomes the lower bound for $O_{ijlm}^{k}(\omega)$, if $x_{ijlm} = 1$. As we minimize the value of $O_{ijlm}^{k}(\omega)$ and $V_{ijm}^{k}(\omega)$, in the

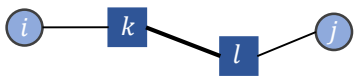
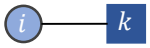
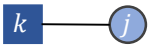
objective function they will attain the lower bound. In line with additional constraints (18-22), we can obtain the linear model.

In proposed model, the routing reliability is calculated by sequentially multiplying the reliabilities of all links and hubs on that route (Eghbali et al. 2014). Since in most real world applications the physical conditions of network facilities are associated to network performance it is essential to consider the efficiency of involved facilities (Kim and O'Kelly 2009).

In this study, reliability factors γ and β are employed to reflect the performance of hub facilities and related connections. Factors γ and β express the degree of benefit from enhancing the reliability of inter hub links and hub facilities when traffic utilizes inter hub links and hub facilities. These factors are set as power parameters (Kim and O'Kelly 2009). Different types of routes in hub location networks and their reliability that are computed by left side of constraint (8). The process of calculating reliability values are shown in Table 2.

Consider that the routing variable X_{iklj} takes one. It means that the traffic from node i to i should pass from three links (ik), (kl) and (jl) with reliabilities Re_{ik} , $Re_{kl}^{1-\gamma}$ and Re_{jl} and two hubs k and l with reliabilities $re_{kk}^{1-\beta}$ and $re_{ll}^{1-\beta}$. Thus, the reliability of this route is calculate by $Re_{ik}re_{kk}^{1-\beta}Re_{kl}^{1-\gamma}re_{ll}^{1-\beta}Re_{jl}$. The reliability of the other potential routes is described in Table 2.

Table 2. Different types of routes in hub location networks and their reliabilities (Eghbali et al. 2014).

Type of routes	Routing variable	Routing reliability
	X_{iklj}	$Re_{ik}re_{kk}^{1-\beta}Re_{kl}^{1-\gamma}re_{ll}^{1-\beta}Re_{jl}$
	X_{ikll}	$Re_{ik}re_{kk}^{1-\beta}Re_{kl}^{1-\gamma}re_{ll}^{1-\beta}$
	X_{ikkj}	$Re_{ik}re_{kk}^{1-\beta}Re_{jk}$
	X_{ikkk}	$Re_{ik}re_{kk}^{1-\beta}$
	X_{kkkj}	$re_{kk}^{1-\beta}Re_{jk}$
	X_{kklj}	$re_{kk}^{1-\beta}Re_{kl}^{1-\gamma}re_{ll}^{1-\beta}Re_{jl}$
	X_{kkll}	$re_{kk}^{1-\beta}Re_{kl}^{1-\gamma}re_{ll}^{1-\beta}$

4. Solution method

In order to solve the problem, proposed mixed-integer linear model is coded by the GAMS 24.1.3. However, only the small-sized instances can be solved to optimality using the general purpose MIP solvers. Due to this limitation, a hybrid meta-heuristic algorithm combined the genetic algorithm and the exact solution approach is used.

Genetic algorithm is a stochastic search approach which imitates the process of evolution and natural selection for finding the near-optimal solution. The primary idea was introduced by Holland (1975). GA works with an initial population of solutions that usually randomly initialized. Each population is represented by chromosomes treat as individuals, whose fitness is calculated by the corresponding objective function value. Individuals of a given population goes through some procedure named evolution consisting of cross over and mutation. Cross over and mutation are used to generate population. In cross over, two chromosomes (parents)

are combined (mated) to produce a new chromosome (offspring). The mutation operator makes random changes in individual's genes. This cycle of evaluation-selection-reproduction is repeated until a well-defined stopping criterion is satisfied. Proposed evolutionary algorithm for solving large-scale problem is composed of meta-heuristic and exact solution based on CPLEX and GA. In fact, our problem can be parsed into two sub-problems: the first stage that design a reliable hub location problem which is an NP-hard problem (Kara and Tansel 2000) and the second stage that specify a service network consists of determining the flight routes and providing booking limits for all itineraries and fare classes after realization of uncertain scenarios. A modified genetic algorithm is employed to find different structures for the reliable hub location network and then to take advantage of the optimal solutions; the CPLEX solver is used in obtaining the overall revenue and other variables based on the network configuration proposed by GA.

4.1. Modified Caching Genetic Algorithm

Caching GA is a modified version of the genetic algorithm with the purpose of avoiding unnecessary calculation of objective values for repetitive individuals during the GA operations (Kratica et al. 2007). Kratica et al. (2007) applied caching techniques to introduce two efficient version of genetic algorithms for solving NP-hard problem. They concluded that implementing caching technique has a significant improvement in the GA running time. They utilized simple and useful caching strategy called Least Recently Used (LRU) with pre-specified data storage.

In the caching process, the objective function value of each individual is stored in a cache table which is called hash-row table. The main advantage of this method is that if we meet the same genetic code during the operation of the genetic algorithm, we can use the stored information in cache-table instead of repeating the calculations. We applied mentioned technique in our proposed algorithm to reduce the time-consuming part of the algorithm which calculates the objective value and sold tickets. By this method, we improve the performance of our evolutionary algorithm significantly. Same as referred work, we used LRU caching strategy and store a certain amount of calculated values in a cache-table with size of $Ncache$=3000. We also considered an additional operator named immigration. In migration operator some immigrants (random members) entrant to the society in each period which keep the algorithm away from local optima during the execution. It is worth mentioning that after obtaining the network structure from the individual's genetic code, by fixing values of w_{ij} and then x_{ijlm} , the initial mixed non-linear integer programming problem is reduced to a linear mixed integer programming (MIP) sub problem that could be solved by CPLEX without the complexity of linearization. According to the mentioned description we used the link between MATLAB software and GAMS. Thereupon, the evolutionary algorithm is coded in MATLAB and a software that makes an interface between MATLAB and GAMS is employed (Ferris 2019).

Complete flowchart of the algorithm is provided in Figure 2. Caching operator is specified with dash line in the figure.

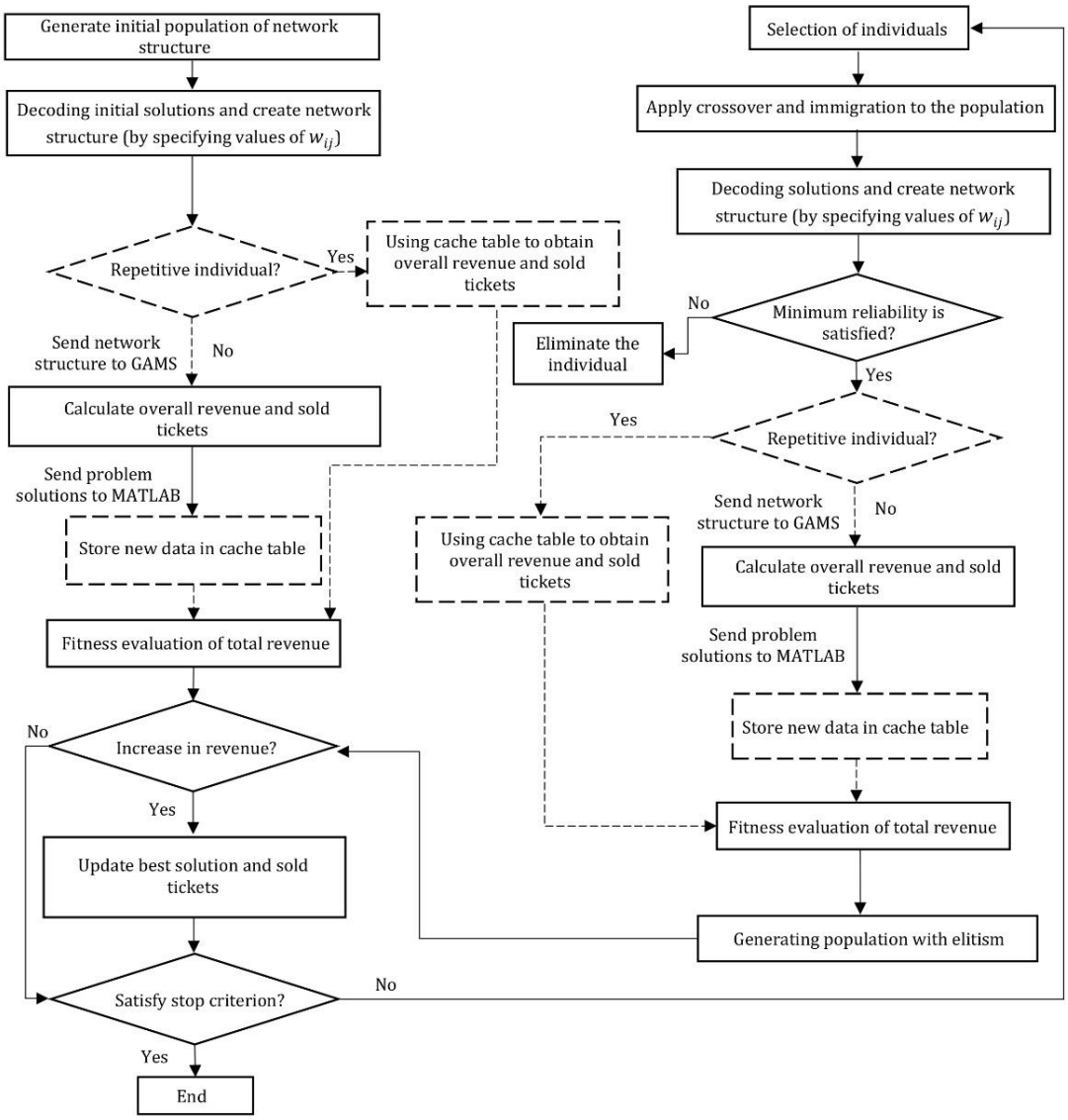


Figure 2. Flowchart of process of proposed evolutionary algorithm based on genetic algorithm

4.2. Representation of the Solution

The efficiency of genetic algorithm is perceptibly affected by chromosome structure. Chromosome structure contains all the information related to solve the problem. The solution of the hub location problem represents the network configuration by demonstrating the location of hubs and allocation of nodes to hubs. We used a ($\boldsymbol{K} \times \boldsymbol{K}$**+1**) dimension matrix contains numbers between [0,1] to represent the given network as a chromosome structure in which $\boldsymbol{K}$ denotes the number of nodes. The first row of the matrix is used to locate the hubs (a ($1 \times \boldsymbol{K}$) matrix). Figure 3 demonstrate an example for location matrix with six nodes and three hubs. In this matrix, the first maximum **P** numbers specify the locations of hubs. This approach ensures that exactly **P** distinct hub indices are established as hubs.

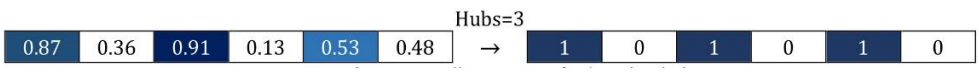


Figure 3. Decoding process for locating hubs.

Rest of the chromosome matrix is utilized to represents the allocation of non-hub nodes to the located hubs. This part of the solution (shown in Figure 4) contains the ($\boldsymbol{K} \times \boldsymbol{K}$) matrix that is filled by random numbers belongs to [0, 1]. The non-hubs nodes are assigned to hubs by comparing the values at the intersection of non-hub node's column and the rows which are assigned as hubs, and the biggest number determines the hub number which the mentioned node should be assigned to that. This approach ensures that each ordinary node is only allowed to connect into one hub. The values which are equal to one at the main diagonal will be considered as a hub and on the other elements with value 1 show the allocated demand nodes. A sample solution related to Figure 4 is depicted in Figure 5.

In the current study we applied two genetic operators consist of crossover and immigration. Thus, the following subsections are dedicated to discuss about these operators.

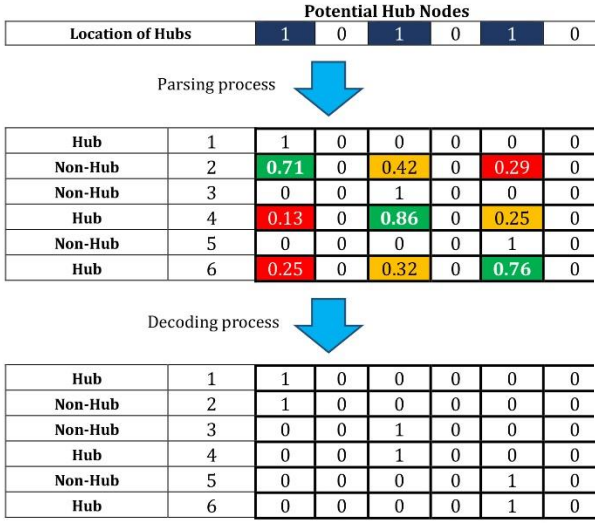


Figure 4. Decoding process for allocation of nodes to hubs.

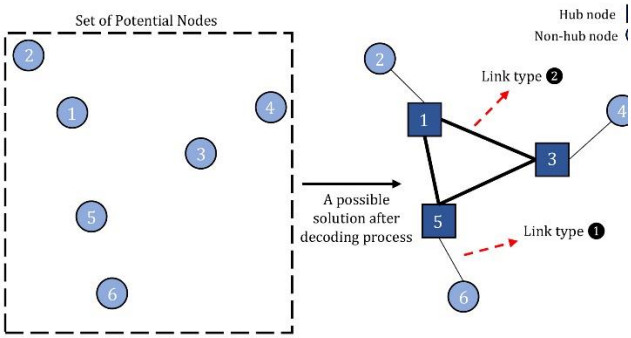


Figure 5. Sample network solution

4.3. Crossover Operator

In crossover operator two individuals are randomly chosen to act as parents so as to create one or more offspring. There are different methods to combine variable values of given parents. In the current study, we applied parameterized uniform crossover. We choose the first parent amongst the best individuals in the population while the other one is chosen from the whole of population, randomly. Then, a real random number in the interval [0,1] for each row is produced. If the random number is larger than a predetermined threshold value, called crossover probability (C_{Prob}), then the allele of the first parent is applied. Otherwise, the allele of the second parent is applied to generate the offspring. An example process of crossover is provided in Figure 6.

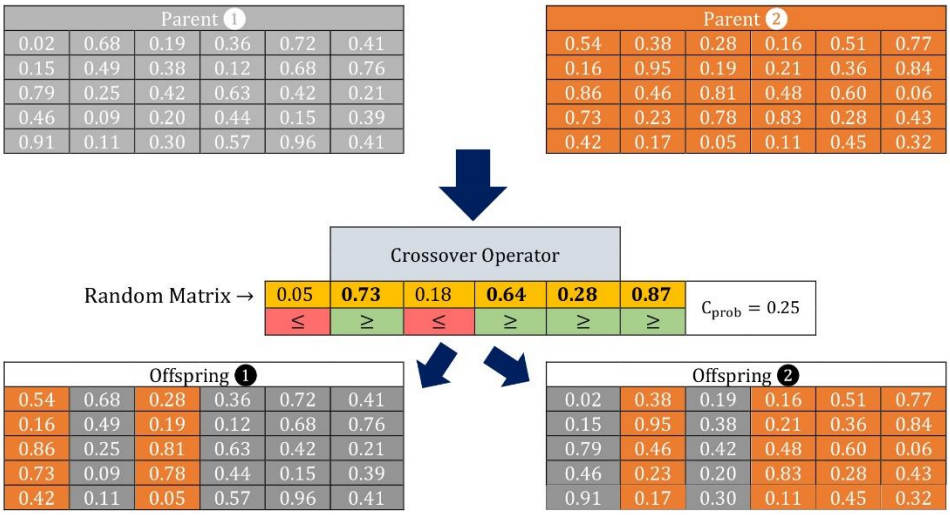


Figure 6. Example of parameterized uniform crossover with crossover probability equal to 0.25.

In this example, the offspring 1 inherits the gene of parent 1 with probability 0.75 and inherits the gene of parent 2 with probability 0.25.

4.4. Migration operator

In this study, instead of using mutation operator we applied immigration operator. Same as mutation operator, immigration operator helps to prevent premature convergence of the population. This idea inspired form many real-world societies in which there is a set of individuals named immigrants which enter to the existing population permanently. These new immigrants are randomly created from the same distribution as the original population and thereupon, no genetic material of the current population is brought in (Tikani and Setak 2019). Altogether, in the proposed algorithm, number of offsprings and immigrants added to the main population, and then after sorting, better individuals enter to the next generation.

5. Computational results

In this section, we reported the results of computational experiments by implementing the model on five subsets of Turkish network data set (Figure 7) presented by Tan and Kara (2007). These subsets are selected from the first 5, 6, 10, 15 and 20 elements of the aforementioned dataset. We also evaluated the performance of implementation of caching technique on modified GA in comparison with the exact solution obtained by GAMS SOLVER 24.1.3 and the performance of the solutions found by the proposed improved genetic algorithm is compared with pure GA and Particle Swarm Optimization (PSO). The meta-heuristic algorithms are implemented in the MATLAB program and linked to the GAMS with ILOG CPLEX 12.5 64-Bit optimization routines. All programs are run on an Intel core i5-3337U (1.8 GHz) with 6 GB of RAM.

To model the uncertainty of data, we applied stochastic programming by using a finite number of scenarios. These scenarios and their relevant probabilities represent an approximation of the probability distribution given by the random data. This method helps to avoid the difficulty of continuous distributions. In the current study, we utilized a scenario tree-based stochastic programming approach to produce demand scenarios. Other input parameters are generated according to Table 3.

Table 3. Generated input parameters

Number of flights between hubs	$DU(3,6)$
Number of flights between nodes and hubs	$DU(1,3)$
The setup costs for a hub	$U(1200000,2400000)$
The unit transportation costs (C)	$U\ [7,14]$
Demand scenarios (high/nominal /low)	125%/100%/75%
Demand classes (economic / business)	80%/20%
Scenario Probabilities	0.19/0.59/0.22
Ticket price for each itinerary for economic class	$U(600,3600)$
Ticket price for each itinerary for business class	$1.5 \times$ ticket price for corresponding economic class
Reliability of links	$U(0.75,1)$
Reliability of hubs	$U(0.80,1)$

In Table 3, $U[a,b]$ denotes a continuous uniform distribution function with upper bound a and lower bound b, and $\mathrm{DU(a,b)}$ denotes discrete uniform distribution with integer parameters between a and b. Parameter C refers to the transfer cost per unit flight and unit distance between origin and hub for business class seats. This is calculated by equation $\mathrm{C}_{11} = \mathrm{C}$, and for economic seats $\mathrm{C}_{12} = \rho\mathrm{C}$ where $\rho \sim \mathrm{U}\ [0.5,1]$. We sampled the nominal demand from Turkish network data set. Then, we considered deviations from the nominal scenario as 125%, and 75% in response to the seasonal demand variation. The capacity of aircrafts in leg between hub and nodes have 100 seats capacity while 200 capacity of seats is considered for aircrafts that transfer passengers between hubs.

Figure 7. Locations of demand nodes for TR data set (Merve and Yaman 2016)

Since parameter calibration has a significant effect on the efficiency of meta-heuristic algorithms, in the next section we applied Taguchi design method for setting the parameters.

5.1. Parameters tuning

Taguchi method provides a fractional factorial experiment instead of full factorial experiments in order to reduce the number of experiments. It divides the factors into signal and noise factors. This method strives to minimize the effect of noise for determining the optimal level of the signal factors.

In current problem, the L9 design of the Taguchi method is employed for meta-heuristic algorithms by using the Minitab 16.2 software. The relative percentage deviation (RPD) formulation is used to change objective function values to non-scale data so a lower response level is more desirable. Since quality characteristic is considered as relative percentage deviation, we selected “Smaller is better” type.

Accordingly, the RPD is calculated as follows:

$$RPD = \frac{|Alg_{sol} - Best_{sol}|}{|Best_{sol}|} \times 100 \tag{23}$$

Where, Alg_{sol} is the objective function value obtained by the meta-heuristic algorithms and $Best_{sol}$ is the optimal solution obtained for each instance by the GAMS software.

In order to study the behavior of different parameters of the proposed evolutionary algorithm, we considered four factors involved in our pure genetic algorithm, namely, the crossover percentage, mutation percentage and population size. For proposed algorithm, we used immigration percentage instead of mutation. Moreover, for PSO we considered three factors consist of number of personal learning coefficient, global learning coefficient and swarm size. The considered levels of the parameters are shown in Table 4.

Relative percentage deviation (RPD) for the objective function value is utilized to compare the levels of parameters. The result of implementing Taguchi design method in MINITAB 16.2 is reported in Table 5.

Table 4. Levels of parameters of meta-heuristic algorithms

Algorithm	Parameter	Level 1	Level 2	Level 3
Modified GA	Crossover percentage	0.60	0.70	0.80
	Immigration percentage	0.05	0.10	0.15
	Population size	50	70	90
GA	Crossover percentage	0.60	0.70	0.80
	Mutation percentage	0.10	0.15	0.20
	Population size	50	70	90
PSO	Personal Learning Coefficient (C1)	1	1.5	2
	Global Learning Coefficient (C2)	1	1.5	2
	Swarm Size	70	80	90

Table 5. Tuned parameters for the proposed GA, pure GA and PSO

Parameter (GA)	Modified GA	GA	Parameter	PSO
Crossover percentage	0.60	0.80	Personal Learning Coefficient (C1)	2
Mutation percentage	-	0.15	Global Learning Coefficient (C2)	1.5
Immigration percentage	0.10	-	Swarm Size	70
Population size	70	70		

5.2. Experimental results

In this section, we reported the results of computational experiments to evaluate the effectiveness of modified GA using MATLAB software in comparison with the exact solution obtained using GAMS SOLVER 24.1.3. Table 6 provides the results of proposed algorithm and exact solution for different instances with discount factor $\alpha = 0.2, 0.4$. In the first and second columns, instance's dimensions n and p are given. The optimum solutions are presented in column three if existed otherwise it is marked with dash. The next column shows computational time in GAMS. We run GA algorithms 10 times and then best obtained solution of the executions is presented in the fifth column. The solution quality is evaluated by $agap = \frac{1}{10}\sum_{i=1}^{10} gap_i$ where $gap_i = 100 \times \frac{Best.sol - sol_i}{Best.sol.sol}$, sol_i indicates the best solution which found in the ith execution, and $Best.sol$ is the optimum solution if it is found and otherwise it is the best solution obtained from all GA runs. In the next column standard

deviation of average gap $\sigma = \sqrt{\frac{1}{10}\sum_{i=1}^{10}(gap_i - agap)^2}$ is presented. In Table 6 $eval$ represents the average number of fitness function evaluations while *cache* shows the average percentage of using cache table to obtain fitness function instead of evaluating it. In fact, it could be considered as a measure of time saving during the executions.

We set the total running time 1 hour as a criterion and proposed GA algorithm terminates if the best solution is not changed during 10 numbers of iterations.

Achieved results obviously demonstrate that the proposed algorithm is an effective solution approach for solving the problem. As can be seen from Table 6, the proposed GA reached the optimum solution in small-sized instances. Therefore, for larger instances where exact methods cannot provide optimum solutions in a reasonable amount of time, proposed algorithm can be used instead. Also, a little gap in achieving the near-optimal solution by the proposed algorithm confirms that presented solution algorithm is efficient in finding the problem solutions.

As can be seen from Table 6, in the experiments with 5 and 6 nodes, exact solution performs faster than the modified hybrid GA. Nevertheless, the computational time in exact method has an exponential growth by increasing the problem size. According to Table 6, changes in discount factor α does not have a meaningful effect on the computational time, however the objective function decreases with increasing the discount factor in all instances. Since we assume different demand levels in various scenarios, the proposed model may provide different flight routes and capacity allocations for various scenarios. For instance, the network structures for two scenarios with 10 nodes and 3 hubs are shown in Figures 8–9.

Proposed model maximizes airline's revenue by designing capacitated network topology and then available capacities among nodes of network partitioned to different customer classes of several itineraries. As we can see from the Figures 8–9, by increasing the demand, the model can earn more profit with fewer itineraries, as the use of hub-stop flights in some routes decreases.

Table 6. Results of optimum solution and proposed GA for different instances

n	p	α=0.2							
		Opt.sol	Time(s)	Modified GA	Time(s)	$eval$	cache	$agap\%$	$\sigma\%$
5	2	578193.19	9.17	opt	31.41	75	0.92	0	0
6	2	1648711.25	40.90	opt	69.96	174	0.89	0	0
10	3	11863082.18	1621.58	opt	371.16	960	0.55	0.46	0.71
	5	21141670.43	1872.19	opt	382.78	874	0.58	0.61	0.92
	7	31409761.71	1923.39	opt	380.18	886	0.60	0.57	0.63
15	3	-	-	16817480.41	578.17	1279	0.48	0.72	1.37
	5	-	-	24429508.20	584.10	1293	0.52	1.01	0.74
	7	-	-	38172945.77	660.27	1248	0.49	0.87	1.38
20	3	-	-	21243181.61	1258.01	1541	0.42	1.93	2.81
	5	-	-	29784726.91	1261.86	1602	0.44	2.02	2.77
	7	-	-	44827554.81	1304.92	1582	0.40	1.95	1.90
		α=0.4							
5	2	561372.29	8.95	opt	33.16	77	0.93	0	0
6	2	1519571.42	41.29	opt	70.14	168	0.90	0	0
10	3	10697139.61	1659.22	opt	354.20	942	0.56	0.47	0.73
	5	20625291.14	1829.98	opt	385.26	869	0.51	0.56	1.11
	7	30847192.26	1966.73	opt	380.02	860	0.58	0.62	0.80
15	3	-	-	15192064.17	581.91	1309	0.46	0.80	0.77
	5	-	-	23573282.14	590.11	1389	0.51	1.04	1.01
	7	-	-	37206711.19	649.44	1360	0.45	0.91	1.22
20	3	-	-	20156594.40	1239.21	1646	0.43	1.76	2.60
	5	-	-	28089941.29	1260.23	1727	0.46	2.12	2.42
	7	-	-	42998018.56	1314.43	1372	0.39	1.88	2.13

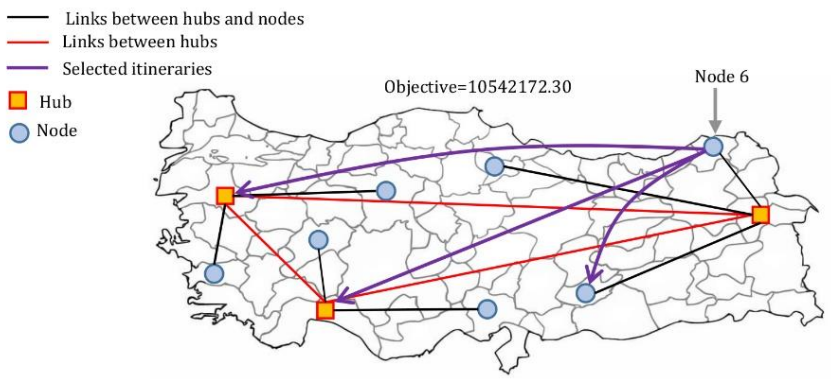


Figure 8. Network and selected itineraries between Node 6 and other cities in high demand.

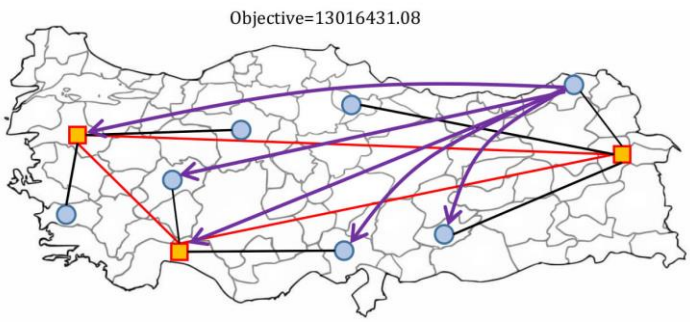


Figure 9. Network and selected itineraries between Node 6 and other cities in low demand.

Table 7. Computational results for the proposed GA, pure GA and PSO

n	p	Modified GA				Pure GA				PSO			
		Obj.value	Time(s)	agap%	σ%	Obj.value	Time(s)	agap%	σ%	Obj.value	Time(s)	agap%	σ%
α=0.2													
5	2	578193.19	31.41	0	0	578193.19	186.16	0	0	578193.19	201.60	0	0
6	2	1648711.25	69.96	0	0	1648711.25	200.08	0	0	1648711.25	228.16	0	0
10	3	11863082.18	371.16	0.46	0.71	11863082.18	715.17	1.35	0.81	11863082.18	712.12	1.01	1.04
	5	21141670.43	382.78	0.61	0.92	21141670.43	694.69	1.25	1.33	21141670.43	702.23	1.30	1.26
	7	31409761.71	380.18	0.57	0.63	31409761.71	712.98	1.90	2.39	31232698.13	730.08	2.17	2.19
15	3	16817480.41	578.17	0.72	1.37	16612391.41	970.29	2.40	2.51	16817480.41	1010.29	2.60	2.67
	5	24429508.20	584.10	1.01	0.74	24429508.20	995. 17	2.57	2.72	24429508.20	1001.26	2.80	2.54
	7	38172945.77	660.27	0.87	1.38	36194422.77	1021.52	3.12	3.48	37569862.18	1049.72	3.11	3.16
20	3	21243181.61	1258.01	1.93	2.81	21243181.61	1539.23	3.87	3.26	21091269.17	1518.82	3.87	3.33
	5	29784726.91	1261.86	2.02	2.77	28956471.40	1719.13	3.99	3.44	27769489.96	1793.17	3.52	3.51
	7	44827554.81	1304.92	1.95	1.90	43948644.20	1984.28	4.14	3.74	43019462.02	2019.28	4.21	3.83
α=0.4													
5	2	561372.29	33.16	0	0	561372.29	188.16	0	0	561372.29	191.19	0	0
6	2	1519571.42	70.14	0	0	1519571.42	196.47	0	0	1519571.42	214.12	0	0
10	3	10697139.61	354.20	0.47	0.73	10697139.61	712.69	1.26	0.90	10697139.61	684.16	1.20	1.02
	5	20625291.14	385.26	0.56	1.11	20625291.14	616.14	1.53	1.11	20625291.14	705.19	1.43	1.39
	7	30847192.26	380.02	0.62	0.80	30847192.26	704.08	2.07	2.17	30847192.26	767.41	2.54	1.94
15	3	15192064.17	581.91	0.80	0.77	15192064.17	1001.11	2.15	1.44	15192064.17	1019.48	2.14	1.56
	5	23573282.14	590.11	1.04	1.01	22065980.01	915.76	3.21	3.01	23573282.14	1015.40	2.62	3.06
	7	37206711.19	649.44	0.91	1.22	36847673.43	1011.19	2.43	3.19	36358521.18	1058.70	2.41	2.79
20	3	20156594.40	1239.21	1.76	2.60	19897429.18	1547.22	4.07	3.42	19968491.35	1489.02	3.67	3.19
	5	28089941.29	1260.23	2.12	2.42	28089941.29	1656.25	4.10	3.86	27576940.40	1786.25	3.56	3.11
	7	42998018.56	1314.43	1.88	2.13	42019816.30	2025.17	3.94	3.88	41767583.09	2004.01	4.50	3.95
Total average :			624.59	0.92	1.23		968.77	2.24	2.12		995.53	2.21	2.07

In order to evaluate the performance of the proposed modified genetic algorithm, we also compared the quality of its solutions with those of solutions obtained by a standard GA algorithm and Particle Swarm Optimization. The results of the comparison for various size instances of the problem are presented in Table 7. We run PSO and GA ten times to solve it. The total computational time and objective function value of the best obtained solution and GAP and standard deviation of ten obtained solutions by PSO and modified GA and pure GA are presented. The results indicate the superiority of proposed algorithm in compared with pure GA and PSO in all problem instances regarding to the quality of the solutions and total computational time.

Figure 10 compares the growth in computational time of the mathematical model with the proposed GA, pure GA and PSO. It is obvious from the figure that by increasing the number of nodes computational time increases in both GAMS and meta-heuristic algorithms. In addition, the proposed caching genetic algorithm has higher speed in compare with ordinary genetic algorithm and PSO due to repetitive individuals that appear during the running process.
Since modified caching genetic algorithm prevents re-calculation of objective function by storing values in a cache table, it can evaluate a large number of population size with higher speed than ordinary genetic algorithm due to repetitive individuals that appear during the GA execution. Figure 11 indicates the speed of objective function improvement in the modified genetic algorithms that tries to achieve the best solution.

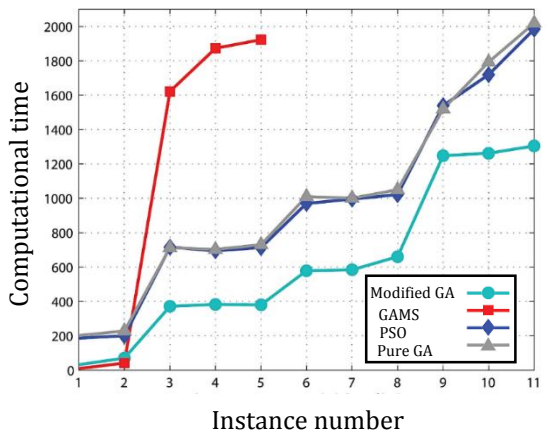


Figure 10. Growth of computational time for different algorithms and GAMS

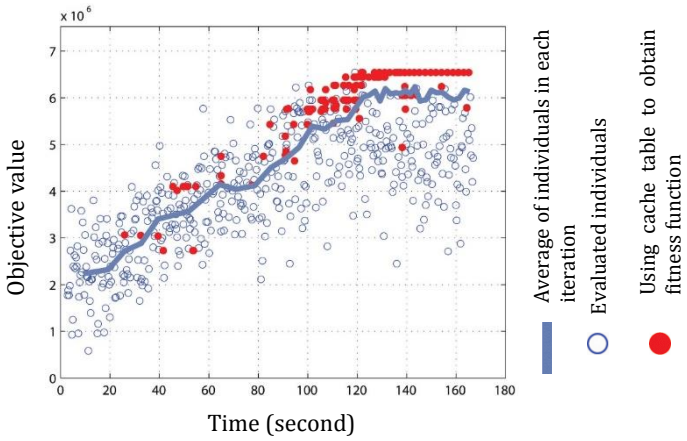


Figure 11. Rate of Convergence in modified GA with problem size 8 node

5.3. Managerial insights

The proposed R-RM-HLP with the profit-oriented objective. In detail, the customers are classified into several groups based on their sensitivity to prices and R-RM-HLP only satisfies a portion of demand nodes which is profitable for the airline company. The following managerial insights can be extracted from this study:

- Our results indicate that applying R-RM-HLP helps to increase the revenue airlines by designing economical capacitated network topology and assign the available capacities to the most profitable flows based on customer classes in various itineraries.
- In designing SC configurations in an uncertain environment. In capacity allocation of R-RM-HLP in an uncertain air traffic passenger, we understand that making decisions

by using stochastic programming yields to increase the flexibility of decisions since the decisions can be varied for different scenarios. The amounts of VSS and $EVPI$ justify the use of complicated formulations and further efforts in the solution techniques.

- The R-RM-HLP would be helpful for decision-makers to consider the entire network's capability in transferring the traffic flows through a hub location problem. In detail, practitioners can compromise between the levels of reliability and the obtained revenue of company. They can control their desire level by adopting the parameter H which shows the reliability threshold for each origin/ destination route.

6. Conclusion

Airline companies encounter with strategic hub location and network design decisions, which has a long lasting impacts on their operations and some tactical flight routes decisions, which are affected by uncertainty. In this study, we focused on demand uncertainty in capacitated reliable transportation network. We formulated and solved the problem as a two stage stochastic programming with the aim of maximizing the airline profit by classifying customers in complete-star transportation network. The proposed model strives to find the optimal locations of p hubs, the allocations of non-hub nodes to the p located hubs and allocate the fixed amount of capacity to various fare classes. Computational study with different number of hubs and nodes was carried out based on Turkish network data set. Since the problem is NP-hard and it is impossible to solve it in large-scales in reasonable amount of time. Therefore, we applied the evolutionary algorithm that includes genetic operators and exact solution by linking MATLAB software to GAMS. In the proposed algorithm, an immigration operator is utilized for a better search in the solution space. This operator helps to preserve the diversity and keep the algorithm away from local optima. In addition, computational performances of the algorithm were improved by caching technique. Caching process prevents re-calculation of fitness function by storing values in a cache table.

The performance of the modified GA is compared with the pure GA and Particle Swarm Optimization (PSO). The results corroborated capability of the proposed algorithm in achieving high-quality solutions in a reasonable time. Moreover, a considerable percentage of run-time savings is obtained by using caching technique. Additionally, the solution quality of proposed genetic search approach is quite satisfactory.

For future works, incorporating pricing strategies, addressing the problem in a competitive market, analyzing the problem with robust optimization and consideration of budget constraints in designing the network structure could be some valuable research fields.

References

Adibi, A. and Razmi, J., (2015). "2-Stage stochastic programming approach for hub location problem under uncertainty: A case study of air network of Iran", *Journal of Air Transport Management*, Vol. 47, pp. 172-178.

Alibeyg, A., Contreras, I. and Fernández, E., (2016). "Hub network design problems with profits", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 96, pp. 40-59.

Alumur, S. and Kara, B.Y., (2008). "Network hub location problems: The state of the art", *European journal of operational research*, Vol.190, No. 1, pp. 1-21.

Alumur, S.A., Nickel, S. and Saldanha-da-Gama, F., (2012). "Hub location under uncertainty", *Transportation Research Part B: Methodological*, Vol. 46, No. 4, pp. 529-543.

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

Aslani, S., Modarres, M. and Sibdari, S., (2014). "On the fairness of airlines' ticket pricing as a result of revenue management techniques", *Journal of Air Transport Management*, Vol. 40, pp. 56-64.

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

Belobaba, P.P. and Weatherford, L.R., (1996). "Comparing decision rules that incorporate customer diversion in perishable asset revenue management situations", *Decision Sciences*, Vol. 27, No. 2, pp. 343-363.

Belobaba, P.P., (1989). "OR practice—application of a probabilistic decision model to airline seat inventory control", *Operations Research*, Vol. 37, No. 2, pp. 183-197.

Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media.

Brumelle, S.L. and McGill, J.I., (1993). "Airline seat allocation with multiple nested fare classes", *Operations research*, Vol. 41, No. 1, pp. 127-137.

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

Çetiner, D., (2013). *Fair revenue sharing mechanisms for strategic passenger airline alliances* (Vol. 668). Springer Science & Business Media.

Chen, S., Gallego, G., Li, M.Z. and Lin, B., (2010). "Optimal seat allocation for two-flight problems with a flexible demand segment", *European Journal of Operational Research*, Vol. 201, No. 3, pp. 897-908.

Cizaire, C. and Belobaba, P., (2013). "Joint optimization of airline pricing and fare class seat allocation", *Journal of Revenue and Pricing Management*, Vol. 12, No. 1, pp. 83-93.

Contreras, I., Cordeau, J.F. and Laporte, G., (2011). "Stochastic uncapacitated hub location", *European Journal of Operational Research*, Vol. 212, No. 3, pp. 518-528.

Čvokić, D.D. and Stanimirović, Z., (2020). A single allocation hub location and pricing problem. *Computational and Applied Mathematics*, Vol. 39, No. 1, pp.1-24.

Damgacioglu, H., Dinler, D., Ozdemirel, N.E. and Iyigun, C., (2015). "A genetic algorithm for the uncapacitated single allocation planar hub location problem", *Computers & Operations Research*, Vol. 62, pp. 224-236.

Davari, S., Zarandi, M.H.F. and Turksen, I.B., (2010), July. The fuzzy reliable hub location problem. In *2010 Annual Meeting of the North American Fuzzy Information Processing Society* (pp. 1-6). IEEE.

Eghbali, M., Abedzadeh, M. and Setak, M., (2014). "Multi-objective reliable hub covering location considering customer convenience using NSGA-II", *International Journal of System Assurance Engineering and Management*, Vol. 5, No. 3, pp. 450-460.

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

Holland, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. U Michigan Press.

Hou, Y.T., Huo, J.Z. and Chu, F., (2019b). "An Integrated Problem of-Hub Location and Revenue Management with Multiple Capacity Levels under Disruptions", *Journal of Advanced Transportation*, 2019.

Huo, J.Z., Hou, Y.T., Chu, F. and He, J.K., (2019a). "A Combined Average-Case and Worst-Case Analysis for an Integrated Hub Location and Revenue Management Problem", *Discrete Dynamics in Nature and Society*.

Kara, B.Y. and Tansel, B.C., (2000). "On the single-assignment p-hub center problem", *European Journal of Operational Research*, Vol. 125, No. 43, pp. 648-655.

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

Kratica, J., Stanimirović, Z., Tošić, D. and Filipović, V., (2007). "Two genetic algorithms for solving the uncapacitated single allocation p-hub median problem", *European Journal of Operational Research*, Vol. 182, No. 1, pp. 15-28.

Lapp, M. and Weatherford, L., (2014). "Airline network revenue management: Considerations for implementation", *Journal of Revenue and Pricing Management*, Vol. 13, No. 2, pp. 83-112.

Lin, C.C. and Lee, S.C., (2018). "Hub network design problem with profit optimization for time-definite LTL freight transportation", *Transportation Research Part E: Logistics and Transportation Review*, Vol. 114, pp. 104-120.

Littlewood, K., (1972). "Forecasting and control of passenger bookings", *Airline Group International Federation of Operational Research Societies Proceedings, 1972*, Vol. 12, pp. 95-117.

Lüer-Villagra, A. and Marianov, V., (2013). "A competitive hub location and pricing problem", *European journal of operational research*, Vol. 231, No. 3, pp. 734-744.

M.C. Ferris (Accessed 2019), MATLAB and GAMS: interfacing optimization and visualization software, University of Wisconsin. http://research.cs.wisc.edu/math-prog/matlab.html.

Meraklı, M. and Yaman, H., (2016). "Robust intermodal hub location under polyhedral demand uncertainty", *Transportation Research Part B: Methodological*, Vol. 86, pp. 66-85.

Momayezi, F., Chaharsooghi, S.K., Sepehri, M.M. and Kashan, A.H., (2018). "The capacitated modular single-allocation hub location problem with possibilities of hubs disruptions: modeling and a solution algorithm", *Operational Research*, pp.1-28.

Mou, D. and Chang, X., (2014). "An uncertain programming for the integrated planning of production and transportation", *Mathematical Problems in Engineering*.

Neamatian Monemi, R., Gelareh, S., Hanafi, S. and Maculan, N., (2017). "A co-opetitive framework for the hub location problems in transportation networks", *Optimization*, Vol. 66, No. 12, pp. 2089-2106.

Nechval, N., Purgailis, M., Rozevskis, U. and Nechval, K., (2013), July. Adaptive Stochastic Airline Seat Inventory Control under Parametric Uncertainty. In *International Conference on Analytical and Stochastic Modeling Techniques and Applications* (pp. 308-323). Springer, Berlin, Heidelberg.

O'kelly, M.E., (1987). "A quadratic integer program for the location of interacting hub facilities", *European journal of operational research*, Vol. 32, No. 3, pp. 393-404.

Setak, M., Feizizadeh, F., Tikani, H., & Ardakani, E. S. (2019). "A bi-level stochastic optimization model for reliable supply chain in competitive environments: Hybridizing exact method and genetic algorithm", *Applied Mathematical Modelling*, Vol. 75, pp. 310-332.

Sim, T., Lowe, T.J. and Thomas, B.W., (2009). "The stochastic p-hub center problem with service-level constraints", *Computers & Operations Research*, Vol. 36, No. 12, pp. 3166-3177.

Taherkhani, G. and Alumur, S.A., (2019). "Profit maximizing hub location problems", *Omega*, Vol. 86, pp. 1-15.

Tan, P.Z. and Kara, B.Y., (2007). "A hub covering model for cargo delivery systems", *Networks: An International Journal*, Vol. 49, No. 1, pp. 28-39.

Tikani, H., and Setak, M. (2019). "Efficient solution algorithms for a time-critical reliable transportation problem in multigraph networks with FIFO property", *Applied Soft Computing*, Vol. 74, pp. 504-528.

Tikani, H., Honarvar, M., and Mehrjerdi, Y. Z. (2016). "Joint optimization of star P-hub median problem and seat inventory control decisions considering a hybrid routing transportation system", *International Journal of Supply and Operations Management*, Vol. 3, No. 3, pp. 1.

Tikani, H., Honarvar, M., and Mehrjerdi, Y. Z. (2018). "Developing an integrated hub location and revenue management model considering multi-classes of customers in the airline industry", *Computational and Applied Mathematics*, Vol. 37, No. 3, pp. 3334-3364.

Topcuoglu, H., Corut, F., Ermis, M. and Yilmaz, G., (2005). "Solving the uncapacitated hub location problem using genetic algorithms", *Computers & Operations Research*, Vol. 32, No. 4, pp. 967-984.

Yang, T.H. and Chiu, T.Y., (2016). "Airline hub-and-spoke system design under stochastic demand and hub congestion", *Journal of Industrial and Production Engineering*, Vol. 33, No. 2, pp. 69-76.

Yang, T.H., (2009). "Stochastic air freight hub location and flight routes planning", *Applied Mathematical Modelling*, Vol. 33, No. 12, pp. 4424-4430.

Yoon, M.G., Lee, H.Y. and Song, Y.S., (2017). "Dynamic pricing & capacity assignment problem with cancellation and mark-up policies in airlines", *Asia Pacific Management Review*, Vol. 22, No. 2, pp. 97-103.

Zarandi, M.H.F., Davari, S. and Sisakht, A.H., (2011), June. Design of a reliable hub-and-spoke network using an interactive fuzzy goal programming. In *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)* (pp. 2955-2959). IEEE.

Zhao, W. and Zheng, Y.S., (2001). "A dynamic model for airline seat allocation with passenger diversion and no-shows", *Transportation Science*, Vol. 35, No. 1, pp. 80-98.

This article can be cited: Salehi, M., Tikani, H., (2020). "Using revenue management technique to allocate the capacity in reliable hub network design under uncertain air passenger traffic", *Journal of Industrial Engineering and Management Studies*, Vol. 7, No. 2, pp. 139-164.