

MATHEMATICAL MODELING OF A BI-OBJECTIVE HUB LOCATION-ROUTING PROBLEM FOR RAPID TRANSIT NETWORKS

MALIHE FALLAH-TAFTI¹, MAHBOOBEH HONARVAR^{1,*}, REZA TAVAKKOLI-MOGHADDAM²
AND AHMAD SADEGHEIH¹

Abstract. This paper aims to develop a mathematical model for rapid transit networks based on a hub and spoke model, comprising stopovers (stations) in the hub and non-hub (spoke) alignments. Due to the use of rapid transit systems in both the hub-level sub-network (*i.e.*, the network among the hub nodes) and the spoke-level sub-network (*i.e.*, the network which connect the spoke nodes to each other and to the hub nodes), the proposed model relaxes some of the usual assumptions in classical hub location models. In the proposed model, the transshipment of flows among the spoke nodes is possible, the setup costs of all the hub and spoke nodes and edges are considerable, and both hub and spoke edges have capacity constraints. In addition to the network infrastructure designed through decisions about the locations of the hub and spoke nodes and edges, the hub and spoke rapid transit lines are determined along with the routes of demands in those lines. The model incorporates profit and service time criteria. An adaptive large neighborhood search solution algorithm is developed whose efficiency is proved by the computational results. Some managerial insight is also provided through the analysis of the resulting networks under various parameter settings.

Mathematics Subject Classification. 90B06.

Received July 26, 2022. Accepted September 28, 2022.

1. INTRODUCTION

Hub location problems (HLPs) focus on the location of hub facilities and the design of hub networks. Hub facilities serve as consolidation or transshipment nodes on origin-destination routes. Transportation systems frequently employ hub and spoke architectures efficiently to route flows between many origins and destinations. Hub nodes provide the possibility of connecting a large number of origin-destination pairs by using a small number of links and the possibility of using economies of scale in transportation costs, as a result of the consolidation of demand flows. In order to benefit from the economies of scale, hub facilities can be connected with highly efficient pathways.

In recent decades, there has been an increasing interest in rapid means of transit such as rapid train, subway and bus rapid transit (BRT) in urban public transportation systems. The use of these rapid public vehicles can

Keywords. Hub location, hub and spoke network, rapid transit network, line planning, adaptive large neighborhood search, bi-objective optimization.

¹ Department of Industrial Engineering, Yazd University, Yazd, Iran.

² School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran.

*Corresponding author: mhonarvar@yazd.ac.ir

be considered as a solution to relieving the congestion that has beset roads and airports in many parts of the world. The establishment of rapid transit systems requires large investment to install stations and links among them. The success of such investment strongly depends on how well those systems are demanded by the public, which is, in turn, dependent on the network design such as the location of stations. With regard to the features of hub models and the large investment needed to build connections in rapid transit systems, using hub structures for designing these systems seems to be advantageous. Some studies, such as [22,23] have proposed pre-assigned topological configurations, *e.g.*, a star, a triangle or a cartwheel, for designing rapid transit networks (RTNs) which have similar structures as hub networks.

The present study aims to develop a mathematical model of designing a rapid urban public transportation network based on a hub-and-spoke model. Instead of using pre-assigned configurations, the network is planned to be based on a general hub structure with stopovers (stations) in the hub and non-hub (spoke) alignments. In fact, what is sought is a hub location model which uses rapid transit systems in both hub-level and spoke-level sub-networks. In the hub-level sub-network, more efficient (larger and faster) vehicles are used to benefit from the economies of scale. For example, in a train network, the hub-level sub-network can be serviced by express trains with high capacity and speed, and the spoke-level sub-network can be serviced by local trains with lower capacity and speed. In addition, different modes of rapid transportation can be considered in the system, *e.g.*, a subway in the hub-level sub-network and BRT in the spoke-level sub-network.

Due to the employment of rapid transit systems in the hub-level and spoke-level sub-networks, the proposed model relaxes some of the common assumptions and properties in classical hub location models. In this regard, direct connections are allowed between spoke nodes, hub and spoke nodes and edges have considerable (nonzero) setup costs, both hub and spoke edges have capacity constraints, the hub-level sub-network is not necessarily a complete network, and paths between origin-destination (OD) pairs do not necessarily contain at least one and at most two hubs. One of the important characteristics of the model is that, unlike generic hub models, it provides possibility not only for direct links between spoke nodes but also for the transshipment of flows at spoke nodes. Given that rapid vehicles usually rout in lines, in the hub network topology, both hub-level and spoke-level sub-networks are considered to be composed of multiple lines. In addition to designing the network, which involves decisions on the location of hub and spoke nodes and the selection of hub and spoke edges, the goal of the research is to simultaneously determine hub and spoke rapid transit lines and the way of routing OD demand flows through these lines.

It should be noted that the proposed model possesses the key distinct features of HLPs, including (a) service demand is associated with flows between OD pairs, (b) hub facilities are intermediate nodes on the OD paths which serve as transshipment or consolidation points, (c) there is the benefit of routing flows *via* hubs; as a requirement, it holds true in situations where the origin and the destination are not located in the same spoke line, and (d) there is a cost-based or service-based objective that depends on the design of the hub network (namely the location of hubs and the selection of links) and the routing of flows [13].

Classical hub location problems are modeled mainly from the viewpoint of costs. They aim to minimize the total cost of a network to satisfy every demand. However, from a profit point of view, it may be more advantageous not to serve every demand. This issue is of importance especially in designing RTNs where setup costs are considerable. Therefore, this research seeks not only to minimize the costs but also to maximize the profit (*i.e.*, revenues minus costs) by assuming no force to satisfy all the demands. In public transportation systems, in addition to the profit of the system operators (managers), the quality of service should also be enhanced to satisfy the users (passengers). As such, this study addresses the maximization of the profit (from the operator's point of view) and the minimization of the service time (from the user's point of view).

Figure 1 illustrates a hub-based RTN. Hub and spoke nodes are presented with filled squares and circles, and hub and spoke edges are shown with wide and narrow links, respectively. The unfilled circles show the demand centers which are not selected to be serviced. In this study, all the demand centers are considered as hub and spoke candidates. In the network shown by this figure, nodes 11 and 18 are not serviced, nodes 8, 14, 15, 16, 17, 19 and 20 are hub stations, and the remaining nodes are spoke ones. The network consists of two hub lines,

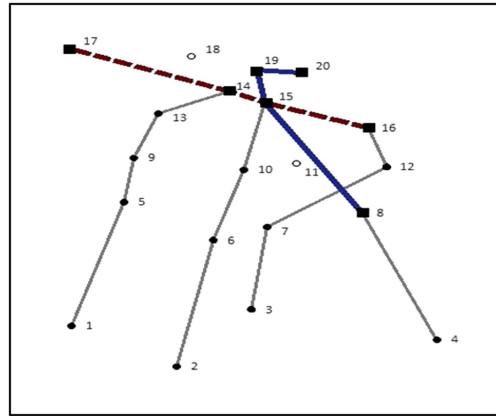


FIGURE 1. A hub-based RTN.

determined by line paths along stations (17, 14, 15, 16) and (8, 15, 19, 20) and four spoke lines by the line paths along stations (1, 5, 9, 13, 14), (2, 6, 10, 15), (3, 7, 12, 16) and (4, 8).

An adaptive large neighborhood search (ALNS) algorithm is proposed to solve the developed model. The model is tested and evaluated on instances derived from two well-known datasets for hub problems. The performance of ALNS is compared with that of the CPLEX solver on small-size instances. The statistical results have proved the efficiency of the proposed algorithm. The resulting hub networks are analyzed under various parameter settings and some managerial insights are given.

The RTN design model proposed based on hubs and spokes has potential applications in designing new RTNs such as subways, BRT systems and multimodal systems in large cities, as in developing countries. It can also be used to improve the existing systems, serve as a decision support tool to evaluate the prospective locations of hub and spoke stations and links, and decide on optimal shipment strategies by considering the criteria for the corresponding operators and users.

The main contributions of this study are as follows:

- A mathematical model is introduced to design a rapid transit network based on a hub and spoke model with stopovers (stations) in the hub and spoke alignments. Rapid transit systems are employed in the hub-level and spoke-level sub-networks.
- The hub-based rapid transit network is designed by decisions about the location of both hub and spoke nodes and the selection of both hub and spoke edges simultaneously. In this regard, the capacity constraints in the hub and spoke edges are taken into account. The model relaxes some of the common assumptions and features of classical hub location models, as the transshipment of flows in the spoke nodes is possible, and the setup costs of all the hub and spoke nodes and edges are considerable (non-zero).
- In addition to the decisions on the network design, decisions are made to determine the hub and spoke lines and the routing of flows through them. From the viewpoint of the hub network topology, this study is based on multiple lines topology for both hub-level and spoke-level sub-networks.
- The criteria for the operators and users are incorporated, *i.e.*, the maximization of profit (revenues minus costs) with no force assumed to service all the demand nodes and the minimization of transit times.
- An ALNS algorithm is developed to tackle the computational complexity of the problem for real-size instances.

The remainder of the paper is organized as follows. The next section presents a survey of the relevant literature. In Section 3, the problem is defined, and a mixed-integer linear programming model is formulated for it. The ALSN algorithm is presented in Section 4, and the results of the computational experiments are reported in Section 5. Finally, Section 6 provides some concluding remarks.

2. LITERATURE REVIEW

This section presents a review of the relevant literature in the two contexts of hub location and rapid transit network design problems.

Five common assumptions underlie most HLPs. They include (a) direct connections between non-hub nodes are not allowed, (b) the costs of edges satisfy the triangle inequality, (c) the discount factor α is constant for hub edges, (d) hub edges have no setup cost, and (e) spoke edges have no setup cost. These five assumptions, without any other restrictions, imply two properties in generic HLPs. First of all, all hubs are fully interconnected by the hub edges, *i.e.*, a hub-level sub-network is a complete network. Secondly, the flows between OD pairs visit at least one and at most two hubs in a hub network in order to minimize the transportation costs. In new approaches, researchers have tried to relax these assumptions in order to make their models realistic enough to be employed in real-world problems. The relaxation of completeness of hub-level sub-networks has attracted significant attention in recent years; it has not been the case for many real-life networks. For example, in the case of transportation networks where the construction cost of edges is not negligible, the completeness of a hub-level sub-network is not based on reality. In some incomplete hub models, no particular topological structure is considered for the hub-level sub-network, not even being connected [7, 8], while some other models have particular structures including circle [15], tree [14], star [31] and simple paths or lines [40, 41]. As spoke-level sub-network can also have these structures or other particular structures such as direct connections [3, 37], multi-stops [30, 57], complete sub-graphs [52] and tours [4, 16, 25, 26, 28, 29, 42]. The choice of the structure depends on the field of application. For a review of the particular topological structures of HLPs, the reader is referred to [13, 30]. As already mentioned, due to employing rapid transit systems in the hub-level and spoke-level sub-networks, the model in this study relaxes Assumptions 1, 4, 5 and, thus, the corresponding properties. From the perspective of the hub network topology and to the best of our knowledge, this study is the first one that uses a topology of multiple lines for both hub-level and spoke-level sub-networks.

Initially, hub networks were mainly used in air transportation. In recent years, considerable attention has been paid to hub structures for designing public transportation networks, which usually are multi-modal, especially those with rapid transit modes. Some of the studies in this case are [12, 24, 27, 36, 40, 41, 43, 51, 54]. In these studies, rapid transit modes are usually considered for the hub-level sub-networks, and the spoke-level sub-networks are usually assigned to road vehicles (such as cars and trucks) [40, 41], buses [24, 43], or ships [21]. Like most of these studies, the present research deals with setup costs for hub edges and, thus, incomplete hub-level sub-networks. However, contrary to the existing hub studies and their application in public transportation, to the best of our knowledge, the present study is the first one that decides on the location of hub and spoke nodes and edges simultaneously. It also determines vehicles transit lines in hub-level and spoke-level sub-networks with capacity constraints in spoke and hub edges, considering both operator and user criteria.

There are not many studies in the literature focusing on maximization objectives in hub location problems especially in non-competitive situations (see *e.g.*, [2, 27, 44, 49, 53]). Like studies such as [2, 44, 49, 53], this research focuses on profit maximization with no concern about all the demand nodes to be necessarily served.

As already mentioned, in recent decades, there has been an increasing interest in rapid transit systems. An RTN design includes two intertwined problems, namely the determination of alignments and the location of stations [13, 33]. In the recent literature on RTN design problems, different aspects of the problem have been captured by optimization models and solved by operations research techniques. In this context, initial efforts were oriented toward determining a single alignment and locating stations given an alignment [32, 46]. In recent years, however, more realistic cases have been investigated. Some of them, such as [22, 23], have dealt with pre-assigned topological configurations, *e.g.*, a star, a triangle and a cartwheel, to design a network with a single mode of rapid transportation and determined optimal rapid transit lines through pre-assigned corridors. Some others such as [20, 34, 35, 38, 39, 48] have investigated RTN designs by determining rapid transit lines, without considering pre-assigned corridors. Tactical decisions have also been incorporated to strategic designs, and integrated RTN design and line planning have been addressed in some problems [9, 10, 17]. Instead of using pre-defined configurations, the present study designs an RTN based on hub and spoke structures with the possibility

of stops (stations) for vehicles in the hub and spoke alignments. In addition to the RTN design, rapid lines are determined in the hub-level and spoke-level sub-networks which constitute the first phase of the line planning process. In RTN design studies, both operator-based and user-based criteria are taken into consideration. The objective functions commonly used in these studies include minimizing costs [5,38,56], minimizing transit times [5,6,19], and maximizing covered demands [20,32,38]. This research also includes both operator-based and user-based criteria by maximizing the total net profit (revenues minus costs) and minimizing the total weighted travel time.

A summary of the main features of the most related studies conducted in the fields of hub location for public transportation networks and RTN design is presented in Table 1. For comparison, the last row of the table presents the characteristics of this study conducted on hub location for RTN design.

3. MODEL AND FORMULATION

In this section, a framework is developed to model the problem of designing a rapid transit network based on hub location. The problem is called a hub-based RTN design problem. It is concerned with locating hub and spoke nodes, allocating spoke nodes to hubs, locating hub and spoke edges, constituting hub and spoke transit lines, determining the percentage of OD demands to be served, and determining the way of routing the demand flows through the network. The objectives of the model are maximizing the total net profit and minimizing the total weighted travel time. Certain capacity constraints are also postulated in the hub and spoke edges. The main assumptions of the developed mathematical model are as follows:

- Each installed spoke node is located in one and only one spoke line.
- The flows can be routed through the included lines.
- Each spoke node is allocated to at most one hub node.
- Without loss of generality, there are the same lower and upper bounds for the length of all the lines, and a different upper bound to the number of nodes exists for the hub and spoke lines.
- In order to calculate the objective functions correctly, all the data related to the costs, revenues and demand flows should be scaled in a same time horizon. For example, if the demand flows are defined for a year or a day, the routing costs and revenues are annual or daily.
- Without loss of generality, the underlying or potential graph used as a basis for building RTN is defined by a set of potential nodes to locate stations and all the possible edges among them.

As a result of these assumptions, the flows can change their lines only at the hub nodes. In other words, if the origin and destination of an OD pair are not located in a same line, the corresponding flow has to be routed through the hub-level sub-network. The hub-based RTN design model involves the following sets, parameters and variables.

Sets

N	Set of the potential nodes to locate the (hub and spoke) stations ($N = \{1, \dots, n\}$).
A	Set of the ordered potential edges among the nodes ($A = N \times N$).
E	Set of edges ($E = \{\{i, j\} : i, j \in N, i < j, (i, j) \text{ or } (j, i) \in A\}$).
LH	Set of possible hub lines (for the ease of counting).
LS	Set of possible spoke lines (for the ease of counting).

Parameters

w_{mn}	The demand flow from node m to node n .
FCh_{ij}	The construction cost of hub edge $\{i, j\}$.
FCs_{ij}	The construction cost of spoke edge $\{i, j\}$.
ch_i	The construction cost of hub node i .
cs_i	The construction cost of spoke node i .
c_{ij}	The transportation cost per unit of flow traveling through edge $\{i, j\}$.

TABLE 1. Comparison of this paper with the most related studies in terms of their characteristics.

Source	Incomplete hub level network	Spoke-level network		Decisions							Objectives				Capacity constraints			Allowing unserved demands	Solution method	
		Direct connection of spokes	Transshipment in spokes	Location							Costs	Incomes	Travel time	Other objectives	Appropriate application in the field of transportation	Hub nodes	Hub edges			Spoke edges
				Hub nodes	Hub edges	Spoke nodes	Spoke edges	Allocation	Flow routing	Other decisions										
[43]	✓	✓	✓	✓	✓		✓	✓	✓	✓					Bi-modal public: <i>e.g.</i> , train in hub level and bus in spoke-level network				Heuristic	
[40]	✓			✓	✓			✓	✓				✓		Public (specially RTN or highway in hub-level network)				Benders branch and cut and heuristic	
[54]	✓	✓	✓	✓				✓	✓	✓			✓	✓	Urban bus				Metaheuristic	
[24]		✓	✓	✓				✓	✓	✓			✓	✓	Bi-modal urban: rail in hub-level and bus in spoke-level network				3-phase heuristic	
[12]		✓	✓	✓				✓			✓	✓			Bi-modal (marine in hub-level and rail in spoke-level network)	✓			Heuristic	
[23]						✓	✓		✓		✓		✓	✓	RTN			✓	ε -constraint	
[9]						✓	✓		✓	✓	✓	✓	✓		RTN			✓	Metaheuristic	
[10]						✓	✓		✓	✓	✓	✓	✓		RTN			✓	Metaheuristic	
This paper	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		RTN		✓	✓	✓	Metaheuristic

d_{ij} The length of edge $\{i, j\}$.

α The discount factor for the transportation cost at the hub edges.

r_{mn} The revenue per unit of flow traveling from node m to node n (sum of the ticket price and the government subsidy)

γh The upper bound on the number of lines which can cross any hub edge.

lh^{\max} The maximum number of the possible hub lines ($lh^{\max} = |LH|$).

ls^{\max} The maximum number of the possible spoke lines ($ls^{\max} = |LS|$).

$hMax$ The upper bound on the number of the hub nodes per hub line.

$sMax$ The upper bound on the number of the spoke nodes per spoke line.

$dMax$ The upper bound on the length of each line.

$dMin$ The lower bound on the length of each line.

$Caph$ The maximum allowed flow in the hub edges.

$Caps$ The maximum allowed flow in the spoke edges (the assumption is $Caps < Caph$).

t_i The time of transfer between the lines in node i per unit of flow.

ah The access time to enter a hub line.

as The access time to enter a spoke line.

eh The exit time to leave a hub line.

es The exit time to leave a spoke line.

λh The average speed of trains in the hub lines ($\lambda h = (1 + (1 - \alpha)) \cdot \lambda s$).

- λs The average speed of trains in the spoke lines.
 M A sufficiently large scalar.

Variables

Variables in network designing and line determination

- x_i If node i is selected to be a spoke node, it is equal to 1, otherwise 0.
 x_{iu} If node i is assigned to hub node u , it is equal to 1, otherwise 0 ($x_{uu} = 1$ if node u is selected to be a hub node; otherwise $x_{uu} = 0$).
 yh_{ij} If edge $\{i, j\}$ is selected to be a hub edge, it is equal to 1, otherwise 0.
 ys_{ij} If edge $\{i, j\}$ is selected to be a spoke edge, it is equal to 1, otherwise 0.
 lh_l If line l is included as a hub line, it is equal to 1, otherwise 0.
 ls_l If line l is included as a spoke line, it is equal to 1, otherwise 0.
 x_i^l If node i is selected to be a spoke node of spoke line l , it is equal to 1, otherwise 0.
 x_{ii}^l If node i is selected to be a hub node of line l , it is equal to 1, otherwise 0 (If l is a spoke line, then $x_{ii}^l = 1$ means that line l and its nodes are assigned to hub node i).
 yh_{ij}^l If hub edge $\{i, j\}$ is selected to be a part of hub line l , it is equal to 1, otherwise 0.
 ys_{ij}^l If spoke edge $\{i, j\}$ is selected to be a part of spoke line l , it is equal to 1, otherwise 0.
 v_l If there is at least one hub line other than hub line l included in the network, it is equal to 1, otherwise 0.

Flow routing variables

- fh_{mni}^l The proportion of the flow from m to n pathing through directed hub edge (i, j) in hub line l .
 fs_{mni}^l The proportion of the flow from m to n pathing through directed spoke edge (i, j) in spoke line l .
 f_{mn} The proportion of the flow from m to n serviced by the RTN.
 τ_{mni} The proportion of the flow from m to n changing the line in node i .

3.1. Modeling of the hub-based RTN design problem

Considering the assumptions and notations in the previous section, the bi-objective formulation for the hub-based RTN design problem is as follows:

$$\begin{aligned}
 \max P = & \sum_{m \in N} \sum_{n \in N} r_{mn} \cdot w_{mn} \cdot f_{mn} - \left[\sum_{m \in N} \sum_{n \in N} w_{mn} \sum_{\{i,j\} \in E} \left(\sum_{l \in LH} \alpha \cdot c_{ij} (fh_{mni}^l + fh_{mnj}^l) \right. \right. \\
 & \left. \left. + \sum_{l \in LS} c_{ij} (fs_{mni}^l + fs_{mnj}^l) \right) + \sum_{\{i,j\} \in E} (FCh_{ij} \cdot yh_{ij} + FCS_{ij} \cdot ys_{ij}) + \sum_{i \in N} (ch_i \cdot x_{ii} + cs_i \cdot x_i) \right] \quad (3.1)
 \end{aligned}$$

$$\begin{aligned}
 \min T = & \sum_{m \in N} \sum_{n \in N} w_{mn} \left[\sum_{j \in N} \left(\sum_{l \in LH} (ah \cdot fh_{mnmj}^l + eh \cdot fh_{mnjn}^l) + \sum_{l \in LS} (as \cdot fs_{mnmj}^l + es \cdot fs_{mnjn}^l) \right) \right. \\
 & \left. + \sum_{\{i,j\} \in E} d_{ij} \left(\frac{1}{\lambda h} \sum_{l \in LH} (fh_{mni}^l + fh_{mnj}^l) + \frac{1}{\lambda s} \sum_{l \in LS} (fs_{mni}^l + fs_{mnj}^l) \right) + \sum_{i \in N: i \neq m, n} \tau_{mni} \cdot t_i \right] \quad (3.2)
 \end{aligned}$$

s.t.

Network design and line determination constraints:

$$yh_{ij} + ys_{ij} \leq 1, \quad \{i, j\} \in E \quad (3.3)$$

$$x_i + x_{ii} \leq 1, \quad i \in N \quad (3.4)$$

$$x_{iu} \leq x_i, \quad i, u \in N, i \neq u \quad (3.5)$$

$$x_{iu} \leq x_{uu}, \quad i, u \in N \quad (3.6)$$

$$\sum_{u \in N} x_{iu} = x_i + x_{ii}, \quad i \in N \quad (3.7)$$

$$\sum_{i \in N} x_{ii}^l = ls_l, \quad l \in LS \quad (3.8)$$

$$x_i^l \leq \sum_{j \in N} x_{jj}^l \cdot x_{ij}, \quad i \in N, l \in LS \quad (3.9)$$

$$x_{ii}^l \leq x_{ii}, \quad i \in N, l \in LH \cup LS \quad (3.10)$$

$$\sum_{l \in LS} x_i^l \leq x_i, \quad i \in N \quad (3.11)$$

$$yh_{ij}^l \leq yh_{ij}, \quad \{i, j\} \in E, l \in LH \quad (3.12)$$

$$ys_{ij}^l \leq ys_{ij}, \quad \{i, j\} \in E, l \in LS \quad (3.13)$$

$$ys_{ij}^l \leq x_i^l + x_{ii}^l, \quad \{i, j\} \in E, l \in LS \quad (3.14)$$

$$ys_{ij}^l \leq x_j^l + x_{jj}^l, \quad \{i, j\} \in E, l \in LS \quad (3.15)$$

$$yh_{ij}^l \leq x_{ii}^l, \quad \{i, j\} \in E, l \in LH \quad (3.16)$$

$$yh_{ij}^l \leq x_{jj}^l, \quad \{i, j\} \in E, l \in LH \quad (3.17)$$

$$\sum_{l \in LH} yh_{ij}^l \leq \gamma h \cdot yh_{ij}, \quad \{i, j\} \in E \quad (3.18)$$

$$\sum_{i \in N} x_{ii}^l \leq hMax \cdot lh_l, \quad l \in LH \quad (3.19)$$

$$\sum_{i \in N} x_i^l \leq sMax \cdot ls_l, \quad l \in LS \quad (3.20)$$

$$dMin \cdot lh_l \leq \sum_{\{i, j\} \in E} d_{ij} \cdot yh_{ij}^l \leq dMax \cdot lh_l, \quad l \in LH \quad (3.21)$$

$$dMin \cdot ls_l \leq \sum_{\{i, j\} \in E} d_{ij} \cdot ys_{ij}^l \leq dMax \cdot ls_l, \quad l \in LS \quad (3.22)$$

$$lh_l \leq \sum_{\{i, j\} \in E} yh_{ij}^l, \quad l \in LH \quad (3.23)$$

$$ls_l \leq \sum_{\{i, j\} \in E} ys_{ij}^l, \quad l \in LS \quad (3.24)$$

$$\sum_{j: j > i} yh_{ij}^l + \sum_{j: i > j} yh_{ji}^l \leq 2, \quad i \in N, l \in LH \quad (3.25)$$

$$\sum_{j: j > i} ys_{ij}^l + \sum_{j: i > j} ys_{ji}^l \leq 2, \quad i \in N, l \in LS \quad (3.26)$$

$$lh_l + \sum_{\{i, j\} \in E} yh_{ij}^l = \sum_{i \in N} x_{ii}^l, \quad l \in LH \quad (3.27)$$

$$ls_l + \sum_{\{i, j\} \in E} ys_{ij}^l = \sum_{i \in N} x_i^l + \sum_{i \in N} x_{ii}^l = \sum_{i \in N} (x_{ii}^l + x_i^l), \quad l \in LS \quad (3.28)$$

$$\sum_{\{i, j\} \in E: i, j \in B} yh_{ij}^l \leq |B| - 1, \quad B \subseteq N, |B| \geq 2, l \in LH \quad (3.29)$$

$$\sum_{\{i,j\} \in E: i,j \in B} ys_{ij}^l \leq |B| - 1, \quad B \subseteq N, |B| \geq 2, l \in \text{LS}. \quad (3.30)$$

Connectivity of the network constraints:

$$x_{ii}^l \leq \sum_{j:j>i} ys_{ij}^l + \sum_{j:i>j} ys_{ji}^l \leq x_{ii}^l + M(1 - x_{ii}^l), \quad i \in N, l \in \text{LS} \quad (3.31)$$

$$\sum_{l' \in \text{LH}: l' \neq l} \sum_{i \in N} x_{ii}^l \cdot x_{ii}^{l'} \geq lh_l \cdot \nu_l, \quad l \in \text{LH} \quad (3.32)$$

$$\nu_l \leq \sum_{l' \in \text{LH}: l' \neq l} lh_{l'} \leq M\nu_l, \quad l \in \text{LH}. \quad (3.33)$$

Flow routing constraints:

$$fh_{mnij}^l + fh_{mnji}^l \leq yh_{ij}^l, \quad m, n \in N, \{i, j\} \in E, l \in \text{LH} \quad (3.34)$$

$$fs_{mnij}^l + fs_{mnji}^l \leq ys_{ij}^l, \quad m, n \in N, \{i, j\} \in E, l \in \text{LS} \quad (3.35)$$

$$\sum_{m \in N} \sum_{n \in N} \sum_{l \in \text{LH}} w_{mn} \cdot (fh_{mnij}^l + fh_{mnji}^l) \leq \text{Caph}, \quad \{i, j\} \in E \quad (3.36)$$

$$\sum_{m \in N} \sum_{n \in N} \sum_{l \in \text{LS}} w_{mn} \cdot (fs_{mnij}^l + fs_{mnji}^l) \leq \text{Caps}, \quad \{i, j\} \in E \quad (3.37)$$

$$\sum_{j \in N} \left(\sum_{l \in \text{LH}} (fh_{mnij}^l - fh_{mnji}^l) + \sum_{l \in \text{LS}} (fs_{mnij}^l - fs_{mnji}^l) \right) \begin{cases} = +f_{mn}, & m, n, i \in N : i = m, m \neq n \\ = -f_{mn}, & m, n, i \in N : i = n, m \neq n \\ = 0, & m, n, i \in N : i \neq m, i \neq n \end{cases} \quad (3.38)$$

$$\tau_{mni} \leq f_{mn}, \quad m, n, i \in N : i \neq m \quad (3.39)$$

$$\sum_{j \in N} (fh_{mnij}^l - fh_{mnji}^l) \leq \tau_{mni}, \quad l \in \text{LH}, m, n, i \in N : i \neq m \quad (3.40)$$

$$\sum_{j \in N} (fs_{mnij}^l - fs_{mnji}^l) \leq \tau_{mni}, \quad l \in \text{LS}, m, n, i \in N : i \neq m. \quad (3.41)$$

Bounding and sign constraints:

$$x_i, x_{iu}, yh_{ij}, ys_{ij}, x_{ii}^l \in \{0, 1\}, \quad i, u \in N, \{i, j\} \in E, l \in \text{LH} \cup \text{LS} \quad (3.42)$$

$$lh_l, yh_{ij}^l, \nu_l \in \{0, 1\}, \quad l \in \text{LH}, \{i, j\} \in E \quad (3.43)$$

$$ls_l, x_i^l, ys_{ij}^l \in \{0, 1\}, \quad l \in \text{LS}, i \in N, \{i, j\} \in E \quad (3.44)$$

$$f_{mn}, \tau_{mni} \in (0, 1)(\text{continuous}), \quad m, n, i \in N \quad (3.45)$$

$$fh_{mnij}^l \in (0, 1)(\text{continuous}), \quad l \in \text{LH}, m, n, i, j \in N \quad (3.46)$$

$$fs_{mnji}^l \in (0, 1)(\text{continuous}), \quad l \in \text{LS}, m, n, i, j \in N. \quad (3.47)$$

Objective function (3.1) maximizes the net profit which is calculated by the subtraction of the total costs, including the transportation costs and the installation costs of the hub and spoke nodes and edges, from the total revenue obtained through satisfying the demands. The revenue is considered as the sum of the ticket price and the government subsidy. The objective function (3.2) minimizes the total weighted travel time. Constraints (3.3) and (3.4) guarantee that each installed node and edge can only be a hub or a spoke. Constraints (3.5)

and (3.6) are added to the model in order to define the variable x_{iu} correctly. Constraints (3.7) hold that each installed node is assigned to one and only one hub node. Each hub node is assigned to itself. Constraints (3.8) enforce each included spoke line to be assigned to one and only one hub node. Constraints (3.9) guarantee that spoke lines assigned to a particular hub node can only consist of the spoke nodes which are assigned to that hub node. Constraints (3.10) mean that a hub station is selected to be the station of a line only if it is already built in the network. Constraints (3.11) ensure that a spoke station is selected to be the station of a line only if it is already built in the network and that it cannot be in more than one spoke line (*i.e.*, flows can change their line only at the hub nodes). Constraints (3.12) and (3.13) enforce an edge to be included in a line if it is already built in the network. Constraints (3.14)–(3.17) ensure that an edge is built in the network only if the adjacent nodes are already built. Constraints (3.18) impose an upper bound on the number of the lines that can circulate at any hub edge of the network, *i.e.*, they prevent a concentration of lines at hub edges (constraints (3.11) enforce this bound to be equal to one for the spoke edges). Constraints (3.19) and (3.20) postulate a maximum number of nodes for each line and impose that no node or edge can be part of a non-included line. Constraints (3.21) and (3.22) assign maximum and minimum lengths to each line. According to the constraints (3.23) and (3.24), at least one edge must be present in each line. Constraints (3.25)–(3.30) determine the lines topology. Constraints (3.25) and (3.26) enforce each node of a line to have at most two incident edges. Constraints (3.27) and (3.28) make the number of the edges of a line equal to the number of its nodes minus one. Constraints (3.29) and (3.30) are sub-tour elimination constraints which guarantee that no line contains cycles. Constraints (3.31) guarantee that the hub node corresponding to a specific spoke line is a node of that line and the line is connected to the hub-level sub-network through that hub node. In fact, these constraints ensure the connectivity of the spoke-level sub-network to the hub-level sub-network and make sure that all the spoke lines are connected to the hub-level sub-network through their corresponding hub nodes. According to constraints (3.25) and (3.26), one can consider $M = 2$ for the constraints already enumerated. Constraints (3.32) postulate connectivity for the hub-level sub-network and mean that each hub line must share at least one node with at least one other included hub line. Constraints (3.33) are added to the model in order to define the ν_l variables correctly. Constraints (3.34) and (3.35) mean that, if the flow corresponding to an OD pair uses an edge of a line, the edge must already be selected to be a part of that line. Constraints (3.36) and (3.37) address the capacity constraints of hub and spoke edges. Constraints (3.38) relate to flow conservation and ensure that the flow from m to n leaves node m , arrives at node n , and is accounted for whenever a middle node i is used. Constraints (3.39)–(3.41) are added to define the τ_{mni} variables correctly. In particular, constraints (3.40) and (3.41) force variables τ_{mni} to take value 1, if the flow from m to n changes its line at node k . Finally, constraints (3.42)–(3.47) are bounding and sign constraints.

3.2. Linearization of the model

To linearize constraints (3.9), the binary variables

$$z_{ij}^l = x_{jj}^l \cdot x_{ij}, \quad i, j \in N, l \in \text{LS},$$

are added to the model, and these constraints are replaced with the following ones.

$$x_i^l \leq \sum_{j \in N} z_{ij}^l, \quad i \in N, l \in \text{LS} \quad (3.48)$$

$$z_{ij}^l \leq x_{jj}^l, \quad i, j \in N, l \in \text{LS} \quad (3.49)$$

$$z_{ij}^l \leq x_{ij}, \quad i, j \in N, l \in \text{LS} \quad (3.50)$$

$$x_{jj}^l + x_{ij} - 1 \leq z_{ij}^l, \quad i, j \in N, l \in \text{LS}. \quad (3.51)$$

To linearize constraints (3.32), the binary variables

$$z_i^{ll'} = x_{ii}^l \cdot x_{ii}^{l'}, \quad i \in N, l, l' \in \text{LH} : l' > l$$

$$z^l = \text{lh}_l \cdot \nu_l, \quad l \in \text{LH}$$

are added to the model, and these constraints are replaced with the following ones.

$$\sum_{l' \in \text{LH}: l' > l} \sum_{i \in N} z_i^{ll'} + \sum_{l' \in \text{LH}: l' > l} \sum_{i \in N} z_i^{l'l} \geq z^l, \quad l \in \text{LH} \quad (3.52)$$

$$z_i^{ll'} \leq x_{ii}^l, \quad i \in N, l, l' \in \text{LH} : l' > l \quad (3.53)$$

$$z_i^{ll'} \leq x_{ii}^{l'}, \quad i \in N, l, l' \in \text{LH} : l' > l \quad (3.54)$$

$$x_{ii}^l + x_{ii}^{l'} - 1 \leq z_i^{ll'}, \quad i \in N, l, l' \in \text{LH} : l' > l \quad (3.55)$$

$$z^l \leq \text{lh}_l, \quad l \in \text{LH} \quad (3.56)$$

$$z^l \leq \nu_l, \quad l, l' \in \text{LH} : l' > l \quad (3.57)$$

$$\text{lh}_l + \nu_l - 1 \leq z^l, \quad l \in \text{LH}. \quad (3.58)$$

4. SOLUTION ALGORITHM

This section presents the proposed solution to the hub-based RTN design problem. First, the weighted sum method is used along with normalization in order to transfer the bi-objective problem into a single-objective one. Then, considering two interconnected sub-problems, an ASLN meta-heuristic algorithm is described to solve the single-objective problem.

4.1. The weighted sum method with normalization

With regard to the concavity of the Pareto frontier of objective functions P and $-T$ for different instances, the weighted sum method is used with normalization in order to handle the bi-objective model proposed in this study. This method is a typical way to solve bi-objective problems, and it works by associating a weighting coefficient with each objective function and optimizing the weighted sum of the objectives. Using the weighted sum method with normalization, the bi-objective problem has been transformed into the following single-objective one:

$$\max z = w \frac{(P - P^{\min})}{(P^{\max} - P^{\min})} - (1 - w) \frac{(T - T^{\min})}{(T^{\max} - T^{\min})}$$

s.t.

Constraints (3.3) to (3.47)

where $w \in [0, 1]$ is the weighting factor which denotes the relative importance of P and T , P^{\max} and T^{\min} are the optimal values of the problem with single objective functions P and T , respectively, and P^{\min} and T^{\max} are the values of functions P and T in the optimal solutions for T^{\min} and P^{\max} , respectively.

Since $T^{\min} = P^{\min} = 0$, the problem is equivalent to the following hub-based RTN design problem called HRTND. The HRTND problem is solved by the proposed meta-heuristic algorithm described in the next subsection.

HRTND problem:

$$\max z = w \frac{P}{P^{\max}} - (1 - w) \frac{T}{T^{\max}}$$

s.t.

Constraints (3.3) to (3.47).

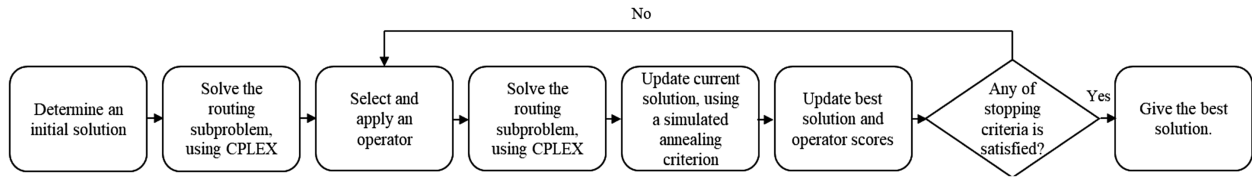


FIGURE 2. Flowchart of the ALNS algorithm.

4.2. An adaptive large neighborhood search method

In order to solve real-size instances of the HRTND problem efficiently, an ALNS meta-heuristic algorithm is proposed here. It is capable of jointly handling the network design, line determination and flow routing decisions. The ALNS meta-heuristic algorithm was first introduced by Ropke and Pisinger [47] in the context of vehicle routing problems and other problems with similar characteristics. It is based on the idea of improving an initial solution iteratively by applying a succession of destroy and repair operators. The success of operators in improving the best found solution enhances their probabilities of being selected. The ALNS algorithm belongs to the category of large-scale neighborhood search techniques defined in [1], but it only examines a relatively low number of solutions. The adaptive part of the algorithm is given by the dynamic updating of the probabilities. The main advantages of the algorithm include the capacity of exploring large parts of the solution space in a structured way, robustness, low probability of getting trapped in local optima, and easy calibration of its parameters [45]. For a survey of ALNS algorithms and applications one can refer to [55]. In the following, the procedure of implementing the proposed ALNS algorithm is explained in detail.

In the presented problem, the ALNS starts with an initial feasible network solution formed by one or more connected lines that satisfy the problem constraints, *i.e.*, relations (3.3)–(3.33) and (3.42)–(3.44). The method of defining this initial network solution will be explained in Section 4.2.2. In each iteration, a new network solution is defined through modifying the current one by applying predefined destroy and repair operators while preserving feasibility. The quality of a network solution is evaluated by solving the routing sub-problem defined with relations (3.1), (3.2), (3.34)–(3.41), and (3.45)–(3.47) where integer variables are fixed according to the network solution, using the CPLEX solver. In fact, given a line network, the flow routing sub-problem solves the problem of determining optimal flow of passengers by maximizing the objective function of the line plan. The new solution is accepted with a simulated annealing criterion by comparing its objective value to that of the current solution. The operators are chosen with a certain probability which depends on their performance in the past iterations, and they are applied on the lines which are selected randomly. The ALNS updates the current and best-known solutions as well as the scores of the last applied operators. The algorithm stops when a stopping criterion (*i.e.*, the maximum number of iterations, the final temperature or the allowed maximum time) is satisfied. These main steps are schematically shown in Figure 2.

In the following, the main components of the ALNS algorithm are described, and then the pseudo-code of the algorithm is presented.

4.2.1. The routing sub-problem

Once a hub RTN is built through the ALNS algorithm, the routing sub-problem is solved to determine the optimal paths of the OD pairs that maximize the weighted sum of the total net profit and the total time and satisfy the capacity constraints. This problem is a linear program and can be solved efficiently by LP solvers such as CPLEX. However, it can actually be solved more efficiently if the number of its variables is reduced. In this case, one can write the variables fs_{mnij}^l in terms of other variables, *i.e.*, the variables fh_{mnij}^l and f_{mn} , as a result of fixing the integer variables and the assumptions mentioned in Section 3. The corresponding formulation includes some notations as defined below:

OD_s The set of OD pairs in which both the origin and destination nodes are located in the same spoke line.

- OD'_s The set of OD pairs minus the set OD_s ($OD'_s = N \times N - OD_s$).
 l_m The spoke line in which spoke node m is located.
 E^{mn} The set of the edges of the spoke line connecting m and n if $m, n \in OD_s$, otherwise an empty set.
 E^m The set of the edges located in part of the spoke line connecting m to the hub-level sub-network if m is a spoke node otherwise an empty set.
 N^m The set of the spoke nodes located in part of the spoke line connecting m to the hub-level sub-network if m is a spoke node, otherwise an empty set.
 $yh_{ij}^{l,g}$ The given values for variables yh_{ij}^l .
 $ys_{ij}^{l,g}$ The given values for variables ys_{ij}^l .

Considering these notations, the routing sub-problem can be redefined as follows:

$$\max w \frac{P}{P_{\max}} - (1 - w) \frac{T}{T_{\max}} \quad (4.1)$$

s.t.

$$\begin{aligned}
 P = & \sum_{m \in N} \sum_{n \in N} r_{mn} \cdot w_{mn} \cdot f_{mn} - \left[\sum_{m \in N} \sum_{n \in N} w_{mn} \sum_{l \in LH} \sum_{\{i,j\} \in E} \alpha c_{ij} \left(fh_{mnij}^l + fh_{mnji}^l \right) \right. \\
 & \left. + \sum_{(m,n) \in OD_s} w_{mn} f_{mn} \sum_{\{i,j\} \in E^{mn}} c_{ij} + \sum_{(m,n) \in OD'_s} w_{mn} f_{mn} \sum_{\{i,j\} \in E^m \cup E^n} c_{ij} \right] \quad (4.2)
 \end{aligned}$$

$$\begin{aligned}
 T = & \sum_{m \in N} \sum_{n \in N} w_{mn} \left(\sum_{i \in N} \sum_{l \in LH} \left(ah \cdot fh_{mnmi}^l + eh \cdot fh_{mnin}^l \right) + \sum_{l \in LH} \sum_{\{i,j\} \in E} \frac{1}{\lambda h} d_{ij} \left(fh_{mnij}^l + fh_{mnji}^l \right) \right. \\
 & \left. + \sum_{i \in N: i \neq m, n} \tau_{mni} \cdot t_i \right) + \sum_{(m,n) \in OD_s} w_{mn} f_{mn} \left(as + es + \frac{1}{\lambda s} \sum_{\{i,j\} \in E^{mn}} d_{ij} \right) \\
 & + \sum_{(m,n) \in OD'_s} w_{mn} f_{mn} \left(as \cdot x_m + es \cdot x_n + \frac{1}{\lambda s} \sum_{\{i,j\} \in E^m \cup E^n} d_{ij} \right) \quad (4.3)
 \end{aligned}$$

$$fh_{mnij}^l + fh_{mnji}^l \leq yh_{ij}^{l,g}, \quad m, n \in N, \{i, j\} \in E, l \in LH \quad (4.4)$$

$$\sum_{m \in N} \sum_{n \in N} \sum_{l \in LH} w_{mn} \cdot \left(fh_{mnij}^l + fh_{mnji}^l \right) \leq \text{Caph}, \quad \{i, j\} \in E \quad (4.5)$$

$$\sum_{m, n \in OD_s: \{i, j\} \in E^{mn}} w_{mn} \cdot f_{mn} + \sum_{m, n \in OD'_s: \{i, j\} \in E^m \cup E^n} w_{mn} \cdot f_{mn} \leq \text{Caps}, \quad \{i, j\} \in E: ys_{ij}^{l,g} = 1 \quad (4.6)$$

$$\sum_{j \in N} \sum_{l \in LH} \left(fh_{mnij}^l - fh_{mnji}^l \right)$$

$$= \begin{cases} \begin{cases} 0, & i \neq m, i \neq n \\ f_{mn}, & i = m, m \neq n, \\ -f_{mn}, & i = n, m \neq n \end{cases} & m, n, i \in N : x_{ii} = 1, x_{mm} = x_{nn} = 1 \\ \begin{cases} 0, & i \neq m, x_{ii}^{l_n} = 0 \\ f_{mn}, & i = m, x_{mm}^{l_n} = 0, \\ -f_{mn}, & x_{ii}^{l_n} = 1, x_{mm}^{l_n} = 0 \end{cases} & m, n, i \in N : x_{ii} = 1, x_{mm} = x_n = 1 \\ \begin{cases} 0, & i \neq n, x_{ii}^{l_m} = 0 \\ f_{mn}, & x_{ii}^{l_m} = 1, x_{nn}^{l_m} = 0, \\ -f_{mn}, & i = n, x_{nn}^{l_m} = 0 \end{cases} & m, n, i \in N : x_{ii} = 1, x_m = x_{nn} = 1 \\ \begin{cases} 0, & x_{ii}^{l_m} = 0, x_{ii}^{l_n} = 0 \\ f_{mn}, & x_{ii}^{l_m} = 1, x_{ii}^{l_n} = 0 \\ -f_{mn}, & x_{ii}^{l_m} = 0, x_{ii}^{l_n} = 1 \end{cases}, & m, n, i \in N : x_{ii} = 1, x_m = x_n = 1, l_m \neq l_n \\ 0, & x_{ii}^{l_m} = 1, x_{ii}^{l_n} = 1 \end{cases} \quad (4.7)$$

$$f_{mn} \leq x_{mm} + x_m, \quad m, n \in N \quad (4.8)$$

$$f_{mn} \leq x_{nn} + x_n, \quad m, n \in N \quad (4.9)$$

$$\tau_{mni} \leq f_{mn}, \quad m, n, i \in N : i \neq m \quad (4.10)$$

$$\tau_{mni} \geq \begin{cases} 0, & m, n \in \text{OD}_s, i \in N, i \neq m, \\ \begin{cases} f_{mn}, & x_{ii}^{l_m} = 1 \text{ or } x_{ii}^{l_n} = 1, \\ 0, & i \in N^m \cup N^n \end{cases}, & m, n \in \text{OD}'_s, i \in N, i \neq m : x_m = 1 \text{ or } x_n = 1 \end{cases} \quad (4.11)$$

$$\sum_{j \in N} (\text{fh}_{mni}^l - \text{fh}_{mnji}^l) \leq \tau_{mni}, \quad l \in \text{LH}, m, n, i \in N : i \neq m \quad (4.12)$$

$$f_{mn}, \text{fh}_{mni}^l, \text{fh}_{mnji}^l, \tau_{mni} \in (0, 1) \text{ (continuous)}. \quad (4.13)$$

Using the above simplified version of the problem reduces the number of variables from $2n^4 + n^2$ to $n^4 + n^2$. As a result, the time required to solve the routing sub-problem reduces considerably, which plays an important role in the efficiency and effectiveness of the proposed ALNS algorithm as this sub-problem is solved in every iteration of the algorithm.

4.2.2. Initial network solution

In this section, a method is proposed to make an initial feasible hub-based RTN solution which consists of several connected hub and spoke lines. This method involves the following three steps.

(1) Determining the location of hubs.

At first, a number p is chosen randomly from the interval $(1, \text{hMax} \times \text{lhMax}]$ as the number of hubs. Then, an efficiency criterion defined as

$$E_i = \frac{\sum_{j=1}^n w_{ij}}{\frac{\text{ch}_i}{\text{cs}_i}},$$

is calculated for each node i , and the p nodes with the highest efficiencies are selected as the hub nodes.

(2) Finding the allocation of nodes.

After selecting p hub nodes, the rest of the nodes are allocated to these hubs. The objective of the optimum allocation model is to minimize the total demand weighted distance. The mathematical framework is as follows.

$$\min \sum_{i \in N'} \sum_{j \in N'} Q_i \cdot d_{ij} \cdot a_{ij}$$

s.t.

$$\sum_{j=1}^p a_{jj} = p \quad (4.14)$$

$$\sum_{j=1}^p a_{ij} = 1, \quad i = 1, \dots, n \quad (4.15)$$

$$b_j \leq \sum_{i \in N': i \neq j} a_{ij} \leq M \cdot b_j, \quad j = 1, \dots, p \quad (4.16)$$

$$\sum_{j=1}^p b_j = \min\{p, \text{lsMax}\} \quad (4.17)$$

where $a_{ij} = 1$ if node i is allocated to hub j ; otherwise, $a_{ij} = 0$. Similarly, $b_j = 1$ if at least one spoke node is allocated to hub j ; otherwise, $b_j = 0$. Moreover, N' is the set N sorted in a descending order; and $Q_i = \sum_{j \in N} (w_{ij} + w_{ji})$. Constraint (4.14) ensures that only p hubs are selected. Constraints (4.15) ensure that each node is allocated to one and only one hub. Constraints (4.16) are needed for defining the variable b_j correctly. In the initial network solution devised in this study, each group of the spoke nodes assigned to a hub together with that hub constitutes a spoke line. Constraint (4.17) sets the number of these groups to the minimum value of p and lsMax.

(3) Generating hub and spoke lines.

The spoke lines are determined through a few steps. At first, among the group of nodes constituting a spoke line, the farthest node from its associated hub is selected as one of the extreme nodes of the line. The other extreme node is the associated hub node. This selection is based on the Euclidean distance. Then, the shortest line passing through the nodes of the group and connecting the defined extreme nodes is considered as the spoke line. After each spoke line is determined, the feasibility of the resulting network solution is checked and the line violating the lower and upper bounds on the length of the line and the upper bound on the number of the nodes of the line is removed from the network. To determine the hub lines, the hub nodes are classified in $[p/\text{hMax}]$ groups. To constitute the first hub line, one of these groups is selected and the two farthest nodes of it are considered as the extreme nodes of the line. Then, the shortest line passing through the nodes of the group and connecting the defined extreme nodes is considered as the first hub line. For determining each of the other hub lines (if needed), one of the remaining groups is selected and its hub nodes together with a hub node randomly selected among the nodes of the last included hub line, are considered as the nodes constituting the line. Among these nodes, the two farthest nodes are considered as the extreme nodes of the line. Then, the shortest line passing through the nodes and connecting the defined extreme nodes, is considered as the hub line. After each hub line is determined, the feasibility of the resulting network solution is checked and the hub line violating the lower and upper bounds on the length of the line and the upper bound on the number of the nodes of hub line are removed from the network. When a hub line is removed, its associated spoke lines are also removed.

4.2.3. Destroy and repair operators

There are six operators involved in the implementation of the presented ALNS. They include two destroy operators, two repair operators and two operators combining both types. The repair operator inserts a new line

or extends an existing one. The destroy operator removes part of a line or a full line. One of the combined operators removes a line and then inserts a new one. The other one removes part of a line and then extends an existing one. In the following, these operators are addressed with more details.

(1) Insert-line operator.

This operator randomly inserts a new line in the network as follows. At first, it randomly decides if it inserts a hub or a spoke line. To insert a hub line, an integer number m is chosen randomly from the interval $[1, \min\{hMax, n_{fh}\}]$, where n_{fh} is the number of the free (non-hub and non-spoke) and the hub nodes of the existing network. Then, m points are randomly chosen among the free and hub nodes, if possible. Finally, considering two farthest nodes as extreme ones, the shortest line passing through the chosen nodes is added to the network if it satisfies the lower and upper bounds on the length of the line, the upper bound on the number of the nodes of the line, not contained in other lines nor containing other lines, upper bound on the number of lines passing through each edge, and the connectivity of the resulting network.

To insert a spoke line, the number m is randomly chosen from the interval $[1, \min\{sMax, n_f\}]$, where n_f is the number of the free nodes. Then, m nodes are randomly chosen among the free nodes as the spoke nodes of the line, if possible. The hub node associated with the line which is one of its extreme nodes is randomly chosen among the hub nodes, and the farthest chosen node from that is considered as the other extreme node. Finally, the shortest line passing through the chosen nodes is added to the network if it satisfies the upper and lower bounds on the length of the line.

(2) Extend-line operator.

This operator extends a randomly chosen line. To extend a hub line, the number m is randomly chosen from the interval $[1, \min\{n_{fh}, hMax - n_s\}]$, where n_s is the number of the nodes of the line. Then, m nodes are randomly chosen among the free and hub nodes, if possible. Finally, the shortest line passing through the chosen nodes and one of the extremes of the existing line, which is chosen randomly, is added to the line if it satisfies the upper bound on the length of the line, the upper bound on the number of the nodes of the line, not containing other lines, and the upper bound on the number of the lines passing through each edge.

To extend a spoke line, the number m is randomly chosen from the interval $[1, \min\{n_f, sMax - n_s\}]$. Then, m nodes are randomly chosen among the free nodes, if possible. Finally, the shortest line passing through the chosen nodes is added to the chosen spoke line if it satisfies the upper bound on the length of the line.

(3) Remove-line operator.

This operator randomly removes an existing line, if the resulting network, after removing the line, remains connected.

(4) Remove-part-line operator.

This operator randomly removes part of a line. First, an existing line is randomly chosen to be partially removed. For the chosen hub lines, the operator removes part of the line between an intermediate node and a terminal one which are randomly chosen. The operator is applied if the new line satisfies the lower bound on the length of the line, not being contained in the other lines, and the connectivity of the resulting network.

For the spoke lines, the operator removes part of the line between an intermediate node randomly chosen and the spoke extreme node. The operator is applied if the new line satisfies the lower bound on the length of the line.

(5) Remove-line and Insert-line operator.

This operator first removes a line using the remove-line operator and then inserts a new line using the insert-line operator, if possible.

(6) Remove-part-line and Extend-line operator.

This operator first removes a part of a line using the remove-part-line operator and then adds a part to a line using the extend-line operator, if possible.

4.2.4. Selection of operators

Let w_i be the weight for operator i , which measures how well the operator has performed recently. The probability of choosing the operator is calculated by $w_i / \sum_j w_j$. In each iteration, the weights are adjusted by means of scores. The scores measure the contribution of the operators to the improvement of the objective function in a block of b iterations.

The score of each operator is initially set to zero in the beginning of the segment. It can be increased at each iteration by σ_1 , σ_2 or σ_3 as follows. If the applied operator results in a new best solution, its score is increased by σ_1 . If it results in a new solution worse than the best solution and better than the incumbent one, its score is increased by σ_2 . If it results in a solution that is worse than the incumbent one but it satisfies the acceptance criterion, its score is increased by σ_3 . In practice, $\sigma_1 \geq \sigma_2 \geq \sigma_3$. After each block of b iterations, all the scores are reset to zero. The weight for operator i can be updated through

$$w_i := \begin{cases} w_i, & o_i = 0 \\ (1-r)w_i + r \frac{s_i}{o_i}, & o_i \neq 0 \end{cases} \quad (4.18)$$

where s_i is the score for operator i , o_i is the number of times that operator i was used in the last segment, and r is the reaction factor parameter which allows controlling how fast the weights change according to the performance in the last block.

4.2.5. Acceptance and stopping criteria

A new solution is accepted according to a simulated annealing acceptance criterion. That is, given the current solution G with an objective value of $z(G)$, the new neighboring network solution G' is accepted if $z(G') > z(G)$. Otherwise, it is accepted with the probability of $e^{(z(G')-z(G))/\eta}$ where $\eta > 0$ is the temperature. The temperature starts with the value $\eta = \eta_0$, but, after a certain number of iterations, it decreases at the cooling rate of $0 < c < 1$. The parameter η_0 is set in a way that the neighboring solution G' , which is at least $\varphi\%$ worse than the current solution, can be accepted with the probability of 50%. The parameter is calculated as follows:

$$\eta_0 = (z(G') - z(G)) / \ln 0.5. \quad (4.19)$$

The execution of the ALNS algorithm ends when a given maximum number of iterations is reached, a given final temperature is attained or the running time exceeds a given threshold.

4.2.6. The ALNS algorithm

In this sub-section, the pseudo-code of the proposed ALSN algorithm is given for more insight (see Algorithm 1).

Algorithm 1. Pseudo-code of the ALNS algorithm.

Input: It consists of the input data of the hub-based RTN model, the initial network solution, and the value of the parameters related to the ALSN algorithm.

Given the initial network solution G_{RTN}^0 , solve the routing sub-problem (formulations (4.1)–(4.13)) to obtain the related solution G^0 and compute the associated objective function $z(G^0)$ using CPLEX.

Set $G = G^b = G^0$ and $z(G) = z(G^b) = z(G^0)$.

While stopping criteria are not satisfied **do**

Select an operator according to $w_i, i = 1, \dots, 6$.

While the selected operator cannot be applied **do**

Select an operator according to $w_i, i = 1, \dots, 6$.

Attempt to apply the selected operator on the current solution G_{RTN} in order to obtain the new network solution G'_{RTN} .

End

Update the number of times the operator is used.

Given G'_{RTN} , solve the routing sub-problem (formulations (4.1)–(4.13)) to obtain the related solution G' and compute the associated objective function $z(G')$ using CPLEX.

If $0 \leq z(G) - z(G') \leq \varphi\%|G|$ for the first time, then adjust η_0 according to (4.19).

If $z(G') > z(G^b)$, then

Set $G^b = G'$, $z(G^b) = z(G')$, and $s_i = s_i + \sigma_1$.

Else, if $z(G') > z(G)$, then

Set $G = G'$, $z(G) = z(G')$, and $s_i = s_i + \sigma_2$.

Else, if G' is accepted by the SA criterion, then

Set $G = G'$, $z(G) = z(G')$, and $s_i = s_i + \sigma_3$.

If the number of the while iterations is a multiple of b , then update the weights of all the operators according to (4.18) and reset their scores to 0.

If $z(G') < z(G^b)$ and η_0 is adjusted, then set $\eta = c \cdot \eta$.

End

Output: It consists of the obtained network infrastructure, a set of lines defined by their itinerary, the values of the satisfied demands as well as their routing, profit, and time.

5. COMPUTATIONAL EXPERIMENTS

This section presents the results of the computational experiments performed to assess the performance of the proposed ALNS algorithm and to analyze the resulting hub RTN networks.

In the first part of the experiments, the ALNS algorithm was compared to the exact branch-and-cut procedure implemented in CPLEX (v12.10), for the HRTND problem. In the second part of the experiments, an analysis was done to check how discount factor and revenue parameters affected the characteristics of the resulting hub-based RTN networks.

All the algorithms were coded in C++ using the Concert Technology of IBM ILOG CPLEX 12.10 to solve linear programs. All runs were implemented on a personal laptop with the Intel®Core™i5-8250U CPU @ 1.60–1.80 GHz processor and Windows 10 with 8 GB of RAM memory.

The computational experiments were conducted with two well-known datasets adopted from the literature of hub location. They included the Australia Post (AP) introduced by Ernst and Krishnamoorthy [18] and the Turkish network dataset which contained the data on travel distances between Turkish cities [50] and the flow between them calculated by the method presented in [11]. The AP dataset provided a symmetric OD demand matrix, while the Turkish network dataset provided an asymmetric one.

In this study, the OD demand and distance matrices provided by the AP and the Turkish network datasets were used, and the values of the other model parameters were set on their basis. As already mentioned, in order to calculate the objective functions correctly, all the data related to costs, revenues and demand flows had to be considered on the same scale. Therefore, the objective functions were calculated in a time horizon of H years. To this end, the demand matrices of the datasets were used to estimate the related demand flows in H years as formulated below. For the AP dataset, w_{ij} , which denotes the number of passengers from i to j in H years, was calculated as

$$w_{ij} = w_{ij}^{\text{AP}} / \sum_{i,j \in N} w_{ij}^{\text{AP}} \times t^{\text{AP}} \times H,$$

where w_{ij}^{AP} is the demand flow of i to j taken from the AP dataset and t^{AP} is the estimated total number of the yearly trips. For the Turkish network dataset, w_{ij} was calculated as

$$w_{ij} = w_{ij}^T \times t^{\text{AP}} \times H,$$

where w_{ij}^T is the demand flow of i to j taken from the Turkish dataset. This parameter is equal to $P_i \cdot P_j / (\sum_{k=1}^n P_k - P_i)$, where P_i is the population of node i , according to [11]. Also t^T is the estimated average number of the yearly trips for each person. The distance matrix of the model was considered equal to the

TABLE 2. Values of the input parameters for the hub-based RTN model.

Name	Description	Value *
n	Number of nodes	$\{10, 20, 30\}$
FCh_{ij}	Construction cost of hub edge $\{i, j\}$ (cur)	$10\,000\,000v_{Fh}d_{ij}, v_{Fh} \in \{8, 10, 12\}$
FCs_{ij}	Construction cost of spoke edge $\{i, j\}$ (cur)	$10\,000\,000v_{Fs}d_{ij}, v_{Fs} \in \{5, 6, 7\}$
ch_i	Construction cost of hub node i (cur)	$10\,000\,000v_{ch}, v_{ch} \in \{23, 25, 27\}$
cs_i	Construction cost of spoke node i (cur)	$10\,000\,000v_{cs}, v_{cs} \in \{8, 10, 12\}$
c_{ij}	Transportation cost per unit of flow traveling edge $\{i, j\}$ (cur/passenger)	$10v_c d_{ij}, v_c \in \{8, 10, 12\}$
α	Discount factor for the transportation cost in the hub edges	$\{0.5, 0, 7, 0.9\}$
r_{mn}	Revenue per unit of flow traveling from node m to n (cur/passenger)	$10v_r d_{mn}, v_r \in \{18, 20, 22\}$
γ_h	Upper bound on the number of the lines circulating on any hub edge	$\{1, 2\}$
lh^{\max}	Maximum number of the possible hub lines	$(1, 2, 3, 4)$
ls^{\max}	Maximum number of the possible spoke lines	$(1, 2, 3, 4)$
$hMax$	Upper bound on the number of hub nodes per hub line	$3, 4, 6$ for $n = 10, 20, 30$, resp
$sMax$	Upper bound on the number of spoke nodes per spoke line	$5, 6, 8$ for $n = 10, 20, 30$, resp
$dMax$	Upper bound on the length of each line	$\max d_{ij}$
$dMin$	Lower bound on the length of each line	$\min d_{ij}$
$Caph$	Maximum allowed flow in the hub edges	$Caps \cdot FCh_{ij}/FCs_{ij}$
$Caps$	Maximum allowed flow in the spoke edges	$v_{Cap} \cdot \max w_{ij}, v_{Cap} \in \{1.8, 2, 2.2\}$
t_i	Time of transfer between the lines in node i per unit of flow (hour)	$v_t \frac{2 \sum_{i,j \in N} d_{ij}/n^2}{(\lambda h + \lambda s)/2}, v_t \in \{0, 0.1, 0.2\}$
ah	Access time to enter a hub line (hour)	$v_a \sum_{i \in N} t_i/n, v_a \in \{0.55, 0.60, 0.65\}$
as	Access time to enter a spoke line (hour)	$ah \cdot \lambda s/\lambda h$
eh	Exit time to leave a hub line (hour)	$v_e \sum_{i \in N} t_i/n, v_e \in \{0.45, 0.50, 0.55\}$
es	Exit time to leave a spoke line (hour)	$eh \cdot \lambda s/\lambda h$
λh	Average of the train speed in the hub lines (hour/km)	$(1 + \alpha_t) \cdot \lambda s$
λs	Average speed of the trains in the spoke lines	80
wp	Weight of the profit objective function	$\{0.25, 0.5, 0.75, 1\}$
M	A sufficiently large scalar.	2
H	Time horizon (year)	30
t^{AP}	Estimated total number of the yearly trips	10 000 000
t^T	Estimated average number of the yearly trips for each person	1

Notes. (*) In cases where a set of values is given for a parameter, the desired value is selected randomly from the set, unless it is explicitly mentioned.

distance matrices provided by the datasets (in terms of km). The remaining parameters were set as reported in Table 2. The values in this table were used in all the experiments, unless it is explicitly mentioned.

For the ALNS algorithm, the score parameters were set to $\sigma_1 = 10$, $\sigma_2 = 5$ and $\sigma_3 = 2$. The cooling rate was set to $c = 0.9997$, the reaction factor to $r = 0.3$, and the number of iterations for updating the weights of the operators to $b = 50$. Furthermore, the initial temperature was set at $T_0 = 10\,000$ at first, and then it was updated during the algorithm implementation considering $\varphi = 0.5$. Finally, the termination criteria were the temperature (T) falling lower than 0.0001, a time limit of 36 000 s (10 h), or a maximum number of 2000 iterations.

TABLE 3. Comparison of the ALNS algorithm with the CPLEX for AP instances.

Row	n	lh^{\max}	ls^{\max}	w^p	CPLEX		ALNS		
					CPU Time	Z	CPU Time	Z	Dev (%)
1	10	1	2	0.25	2	0	113	0	0.0
2				0.50	38	0.211365	135	0.19774	6.4
3				0.75	520	0.596002	133	0.572356	4.0
4				1.00	1279	1	141	0.9136745	8.6
5		2	4	0.25	3.16	0	282	0	0.0
6				0.50	277	0.2251851	318	0.1944262	13.7
7				0.75	7534	0.6075345	350	0.5193543	14.5
8				1.00	35 900	1	374	0.8611032	13.9
9	20	2	4	0.25	59.19	0	2500	0	0.0
10				0.50	36 008	0.2005308	3591	0.1918571	4.3
11				0.75	36 004	0.5190457	4137	0.6018812	−13.8
12				1.00	36 004	0.8014915	3807	1	−19.9
13		3	6	0.25	68.34	0	5280	0	0.0
14				0.50	36 006	0.0006236	6561	0.2246729	−99.7
15				0.75	36 008	0	5750	0.596825	−100.0
16				1.00	36 004	0	5399	1	−100.0
17	30	3	6	0.25	OM	OM	36 002	0	—
18				0.50	OM	OM	35 245	0.2799198	—
19				0.75	OM	OM	36 002	0.5364978	—
20				1.00	OM	OM	36 002	1	—
21		4	8	0.25	OM	OM	36 015	0	—
22				0.50	OM	OM	36 018	0.2564526	—
23				0.75	OM	OM	36 013	0.4786503	—
24				1.00	OM	OM	36 036	1	—

5.1. ALNS algorithm performance

In order to gain an insight into the performance of the ALNS algorithm, its capability to reach an optimal solution was investigated through comparing it to the CPLEX solver.

In this subsection, experiments were conducted with $n \in \{10, 20, 30\}$ for the AP and Turkish network datasets. For each number of n , some instances were solved with two values of the maximum number of the hub and spoke lines, *i.e.*, $(lh^{\max}, ls^{\max}) \in \{(1, 2), (2, 4)\}$ for $n = 10$, $(lh^{\max}, ls^{\max}) \in \{(2, 4), (3, 6)\}$ for $n = 20$, and $(lh^{\max}, ls^{\max}) \in \{(3, 6), (4, 8)\}$ for $n = 30$. In each case, the weight of the profit function was varied with regard to $w^p \in \{0.25, 0.50, 0.75, 1.00\}$, and the discount factor was fixed to 0.7. This experimental framework was dealt with in 24 instances. In all the experiments, a CPU time limit of 10 h was used for the CPLEX solver, as for the ALNS algorithm. In order to better compare the solutions produced by the methods, the parameters P^{\max} and T^{\max} were calculated *via* the CPLEX solver for instances with $n = 10$, where CPLEX could reach the exact optimal solution. For the remaining instances, the parameters were determined *via* the ALNS algorithm, which usually yield better values for the parameters.

Tables 3 and 4 present the computational results of comparing the ALNS method and the CPLEX in terms of the computational time and the quality of the solutions. The comparisons were performed for AP and Turkish network datasets. In order to compare the solutions produced by the two methods, a performance measure was reported in the “Dev (%)” columns. It was the percentage of the deviation between the solutions found by the ALNS method and the CPLEX, considering only the instances for which the CPLEX obtained a feasible solution. This measure was calculated as $100 \times (z^{\text{CPLEX}} - z^{\text{ALNS}}) / \max\{z^{\text{CPLEX}}, z^{\text{ALNS}}\}$, where z^{CPLEX} and

TABLE 4. Comparison of the ALNS algorithm with the CPLEX for AP instances.

Row	n	lh^{\max}	ls^{\max}	wp	CPLEX		ALNS		
					CPU Time	Z	CPU Time	Z	Dev (%)
1	10	1	2	0.25	1.86	0	122	0	0.0
2				0.50	175.7	0.239491	136	0.230569	3.7
3				0.75	851.20	0.5952738	136	0.5491577	7.7
4				1.00	3464.81	1	131	0.9147655	8.5
5		2	4	0.25	2.17	0	282	0	0.0
6				0.50	765.59	0.2264264	192	0.2038717	10.0
7				0.75	2223.69	0.6033136	209	0.5339765	11.5
8				1.00	4341.48	1	322	0.9048438	9.5
9				0.25	53	0	3336	0	0.0
10	20	2	4	0.50	36 000	0.2105771	3223	0.0944445	55.1
11				0.75	36 000	0.6620787	4248	0.7208722	−8.2
12				1.00	36 000	0	4224	1	−100.0
13		3	6	0.25	72	0	4309	0	0.0
14				0.50	36 000	0.0527871	6359	0.1338896	−60.6
15				0.75	36 000	0	5852	0.5432367	−100.0
16				1.00	36 000	0	6623	1	−100.0
17				0.25	OM	OM	36 016	0	—
18	30	3	6	0.50	OM	OM	30 584	0.2866742	—
19				0.75	OM	OM	36 016	0.6065127	—
20				1.00	OM	OM	36 016	1	—
21				0.25	OM	OM	36 016	0.0026458	—
22		4	8	0.50	OM	OM	36 010	0.2463013	—
23				0.75	OM	OM	36 020	0.6593838	—
24				1.00	OM	OM	36 028	1	—

z^{CPLEX} are the objective values obtained by the CPLEX and the ALNS, respectively. In situations in which both objective values are equal to zero, the measure is defined equal to zero. The “CPU Time” columns present the consumed CPU times of the methods in seconds. Also, the “ Z ” columns present the obtained objective values of the methods. Concerning the ALNS execution, three different experiments were conducted for each instance, the average results of which are reported in Tables 3 and 4.

According to the tables, the ALNS outperforms the CPLEX in part as a consequence of the big size of the mixed integer programming model. The advantage of the ALNS increases as the problem becomes more difficult, *i.e.*, when the nodes and/or lines increase in number. Note that the CPLEX failed to solve the instances with 30 nodes for both datasets, which was due to excessive memory requirements. These out-of-memory cases are indicated by “OM”. According to the reported optimality deviations, the CPLEX presented better solutions for the instances with 10 nodes, while the ALNS performed better for the other instances. In more than 93% of the cases, the percent deviation between the two objective values was less than 15%, which proves the efficiency of the ALNS algorithm in solving the HRTND problem. More precisely, as the tables suggest, the ALNS method yielded good solutions with the percent deviation of −16.6% on average. Furthermore, the ALNS algorithm solved the problem faster on average. The superiority of the ASLN over the CPLEX increased as the problem became larger. This behavior was similar for both datasets. It can be concluded that the use of the commercial CPLEX solver is not adequate for solving complex HRTND problems.

Figures 3 and 4 represent a detailed comparison of the processes of convergence to the best solutions obtained by the ALNS and CPLEX methods through the experiments reported in Tables 3 and 4, respectively. The vertical axis represents the running time, whereas the horizontal one corresponds to the value of the objective function.

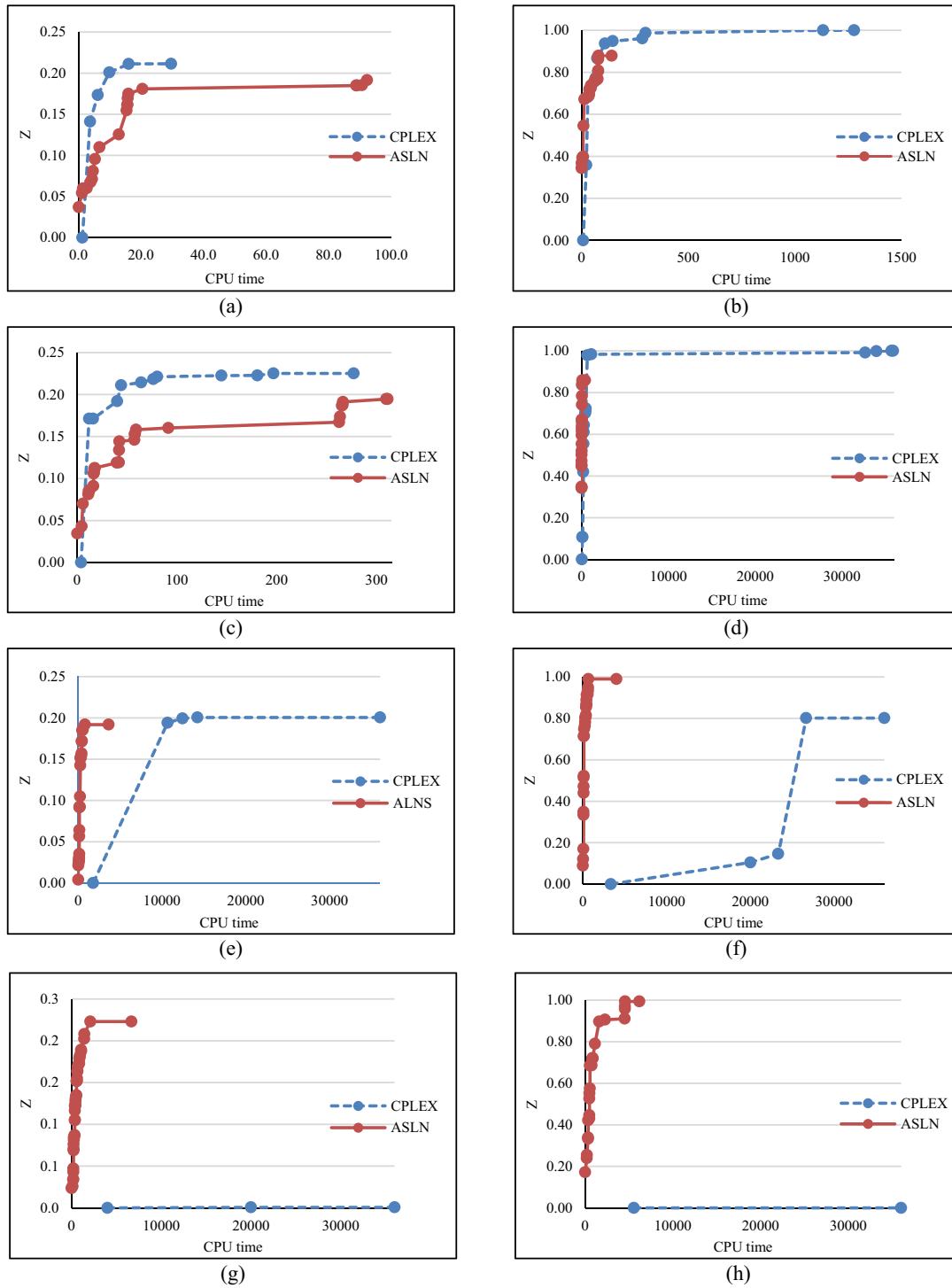


FIGURE 3. Convergence of ALNS and CPLEX for the experiments corresponding to the (a) 2nd, (b) 4th, (c) 6th, (d) 8th, (e) 10th, (f) 12th, (g) 14th, and (h) 16th rows of Table 3 (AP related experiments).

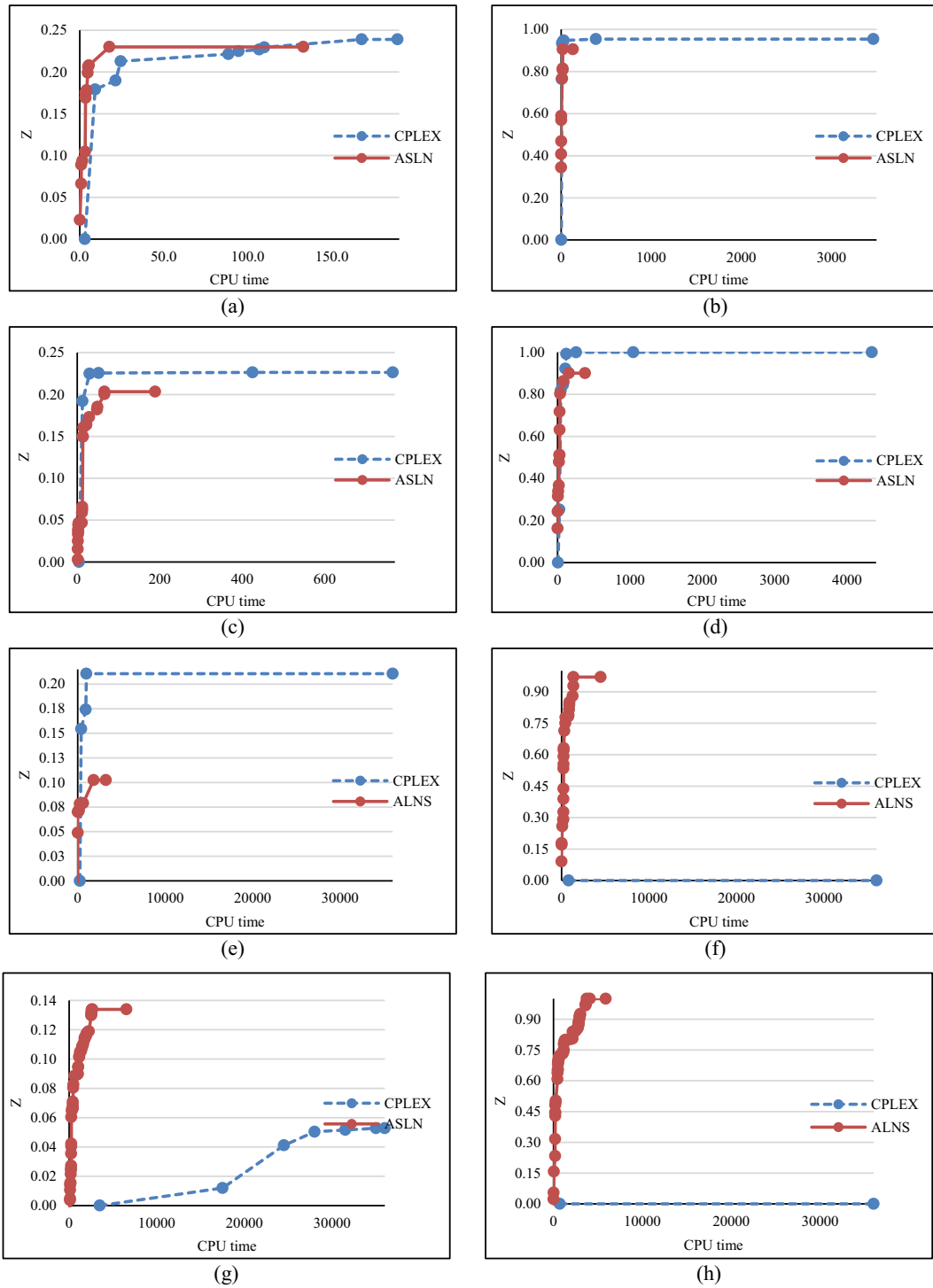


FIGURE 4. Convergence of ALNS and CPLEX for the experiments corresponding to the (a) 2nd, (b) 4th, (c) 6th, (d) 8th, (e) 10th, (f) 12th, (g) 14th, and (h) 16th rows of Table 4 (Turkish network related experiments).

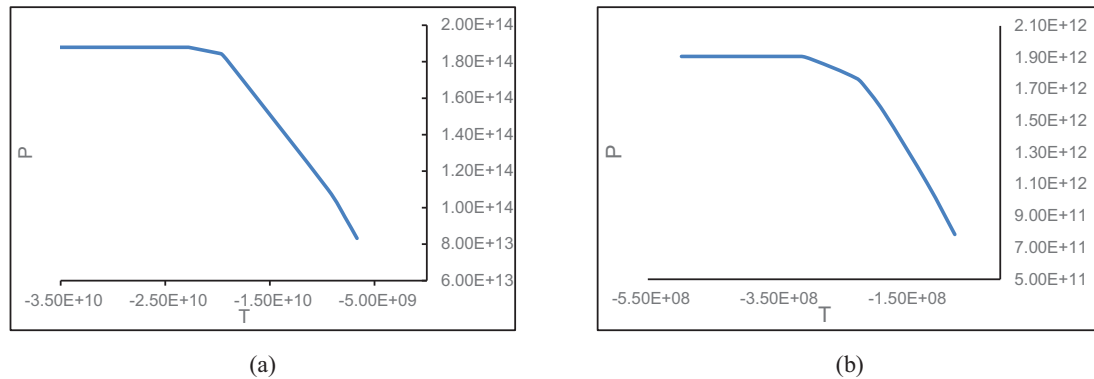


FIGURE 5. Pareto frontier for (a) AP and (b) Turkish network instances.

The results related to the mean objective values obtained in three runs of the ALNS method are presented in the figures. As it can be seen and as already mentioned, the ALNS method could reach good solutions within much less time than the CPLEX. More precisely, the ALNS averagely captured more than 85% of the best objective value obtained by both methods in less than 25% of the maximum CPU time of the methods, which means its convergence speed is very high. The figures also suggest that the ALNS method usually begins with a considerably better initial solution than the CPLEX, which confirms the affectivity of the method used to determine an initial solution.

In a part of the numerical investigation, the Pareto frontier was plotted for some instances by changing the weights of the objective functions, through which the frontier emerged to be concave. This means that using the weighted sum method is reasonable for handling the bi-objectivity of the problem addressed in this study. The Pareto frontiers of the experiments with $n = 10$, $h = 1$ and $s = 2$ for the AP and Turkish network datasets are given in Figure 5 as some examples.

5.2. Analyzing the network design

In the second part of the experiments, analyses were performed of the discount factor, α , and the revenues (sum of the ticket prices and the government subsidies), $r_{mn}s$, as two important parameters which affect the design of hub-based RTN networks and can be determined based on decision makers' policies. To this end, the obtained solutions were investigated mainly in terms of the total net profit and the total travel time values, the percentage of the served demands and the topology of the network. Then, some managerial insights were given.

In this subsection, the three experiments explained in the previous subsection are referred to again. They concern problems with the size of $n = 10$, $lh = 1$, $ls = 2$, $n = 20$, $lh = 2$, $ls = 4$ and $n = 30$, $lh = 3$, $ls = 6$ for the AP dataset with $w^p = 0.5$. These experiments correspond to the second, 10th and 18th rows of Table 3, and their best solution network configurations are presented in Figure 6. The thick lines represent the hub lines (different hub lines are given in different dashed types), and the narrow ones represent the spoke lines. The nodes that are not connected to the network are the ones which are not serviced. For each experiment, some instances were solved with the varied values of the parameter under investigation but the fixed values of the other parameters. In each case, the values of P^{\max} and T^{\max} were the same as those in the experiments of the previous subsection. For the instances with $n = 20$ and $n = 30$, where there was no optimal solution by the exact solver CPLEX, the solution provided by the ALNS algorithm was used. In these cases, the best solutions found in three runs are reported here.

Table 5 gives the results of the experiments conducted to investigate the effect of α . For each size of the problem, the next four columns indicate the corresponding values for the discount ($1 - \alpha$), the total percentage of the satisfied demand, the total net profit, and the total travel time, respectively. The results show that the

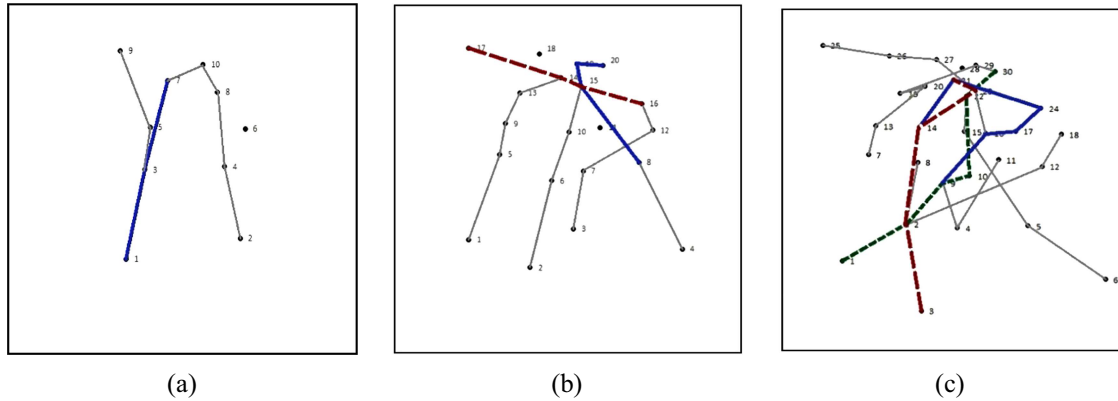


FIGURE 6. Configuration of the networks obtained for the experiments corresponding to the (a) 2nd, (b) 10th, and (c) 18th rows of Table 3.

TABLE 5. Solution characteristics for different values of the discount factor with the AP dataset.

Row	Size	$1 - \alpha$	Satisfied demand (%)	Profit value	Time value
1	$n = 10$,	0.1	38	1.63734e+14	1.9169e+10
2	$lh^{\max} = 1$,	0.3	38	1.84267e+14	1.96267e+10
3	$ls^{\max} = 2$	0.5	38	2.02905e+14	1.96267e+10
4	$n = 20$,	0.1	46	1.97771e+14	2.80878e+10
5	$lh^{\max} = 2$	0.3	40	1.95987e+14	2.52526e+10
6	$ls^{\max} = 4$	0.5	41	2.30737e+14	2.95268e+10
7	$n = 30$,	0.1	33	1.28679e+14	2.05161e+10
8	$lh^{\max} = 3$,	0.3	32	1.74225e+14	2.32456e+10
9	$ls^{\max} = 6$	0.5	34	1.87293e+14	2.38087e+10

decision maker can obtain significantly more profit when the discount is higher due to the economies of scale. The choice of this parameter does not seem to have a decreasing or increasing effect on the value of the total time and the percentage of the captured demand. Figure 7 presents the resulting network configurations. As it can be seen, the ratio of the total length of the hub lines to that of the spoke lines becomes greater or remains the same when the discount $1 - \alpha$ increases. In detail, this ratio is 0.54 for the first three rows of the table, 0.45, 0.58, and 0.67 for rows 4–6, and 0.71, 0.78, and 0.94 for rows 7–9.

Similarly, Table 6 presents the results of the experiments in the investigation of the effect of $r_{mn}s$. For this investigation, the relation $r_{mn} = \gamma d_{mn}$ was taken into account, and the parameter γ was varied as given in Table 6. Figure 8 presents the resulting network configurations. As understood from Table 6 and as expected, an increase in the revenue parameter γ leads to the rise of the percentage of the satisfied demand and the total net profit. It also increases the total time of transportation. This can be because having more satisfied passengers may result in exceeding the capacity of edges on shorter OD paths and the necessity of routing the passengers in longer ones. According to Figure 8, the parameter has a significant impact on the network topology. Lower revenues give rise to shorter networks with lower numbers of stations and edges. This is because as revenues and, therefore, satisfied demands decrease, high investment for building a large number of stations and edges cannot be profitable.

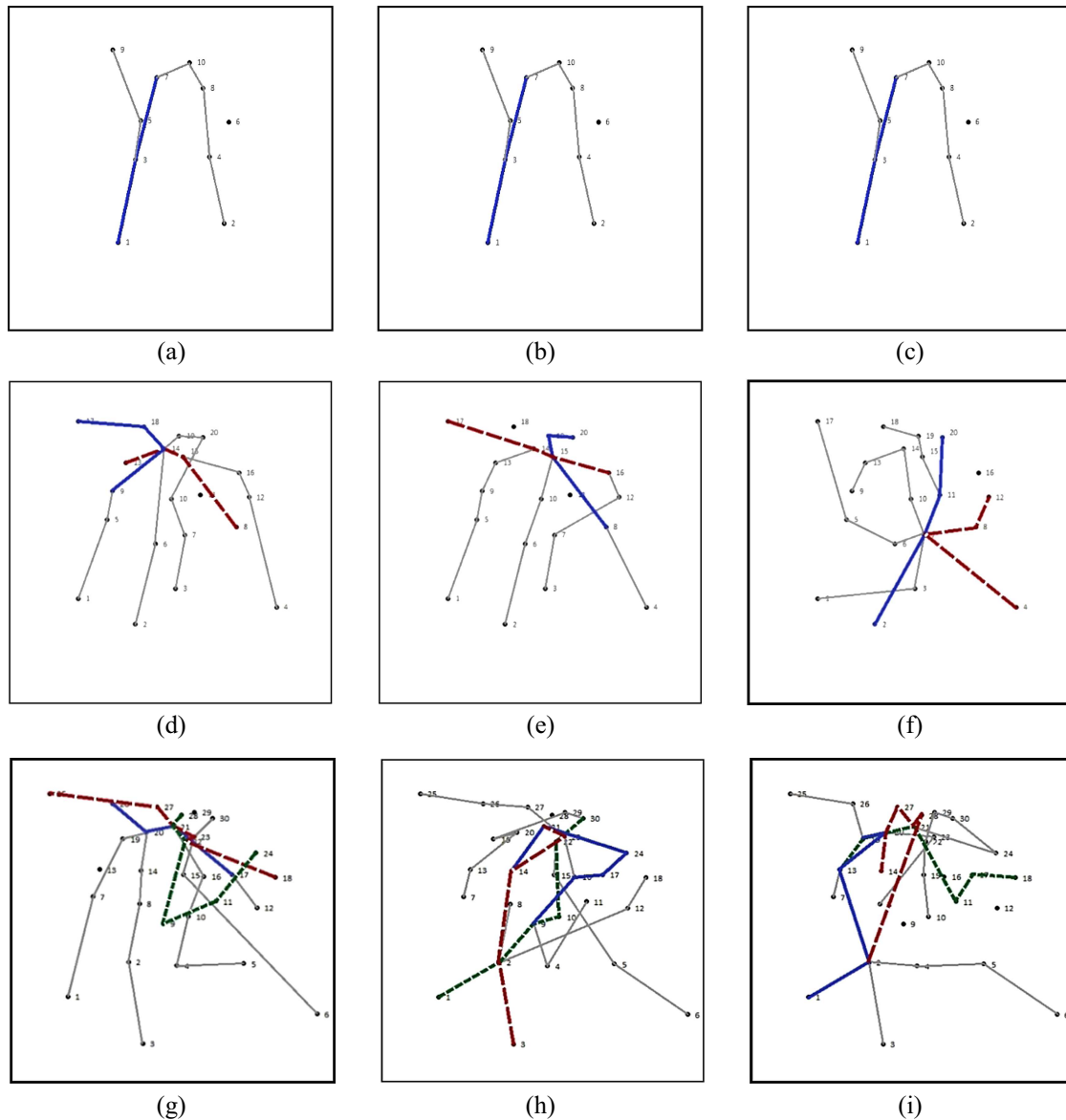


FIGURE 7. Configuration of the networks obtained for the experiments corresponding to the (a) 1st, (b) 2nd, (c) 3rd, (d) 4th, (e) 5th, (f) 6th, (g) 7th, (h) 8th, and (i) 9th rows of Table 5.

Using the computational results, the effectiveness and efficiency of the proposed model were validated, and useful insights were provided into the interactions among the different aspects of the studied complex decision problem. From these results, one can conclude that different planning decisions are jointly involved in obtaining good solutions. This shows the importance of considering different aspects and phases of a planning process in order to obtain good solutions for real-life situations. The proposed model can, thus, be employed to design new rapid transit systems successfully as different decision-making parameters are taken into consideration. The model can also be used as a decision support tool to improve or extend the existing systems through investigating the effects of important parameters on the resulting networks while considering the criteria of net profit and

TABLE 6. Solution characteristics for the different values of the revenue factor with the AP dataset.

Row	Size	γ	Satisfied demand (%)	Profit value	Time value
1	$n = 10,$	150	0.15	6.0912e+13	8.6176e+09
2	$lh^{\max} = 1,$	200	0.38	1.84267e+14	1.96267e+10
3	$ls^{\max} = 2$	250	0.38	2.67534e+14	1.96267e+10
4	$n = 20,$	150	0.21	6.44126e+13	1.10989e+10
5	$lh^{\max} = 2,$	200	0.40	1.95987e+14	2.52526e+10
6	$ls^{\max} = 4$	250	0.37	2.97146e+14	2.61594e+10
7	$n = 30,$	150	0.09	6.12836e+13	1.95175e+10
8	$lh^{\max} = 3,$	200	0.32	1.74225e+14	2.32456e+10
9	$ls^{\max} = 6$	250	0.32	2.56074e+14	2.64467e+10

time. For example, as mentioned, a producer can use more efficient vehicles in hub lines and give a more discount to gain more profit without effecting the total time significantly. In addition, an increase in the revenue (*i.e.*, the sum of the ticket prices and the attracted governmental subsidies) raises both the total net profit and the time. Therefore, one can decide about these parameters based on the degree of the importance of the two criteria. Of course, for more accurate analysis, new tactical decisions (*e.g.*, determining the optimal fleet size considering the fleet acquisition costs) or elastic demands (*e.g.*, price sensitive demands) may be incorporated. These issues are outside the scope of this paper, as the goal is here to propose an efficient basic model of designing hub-based RTNs; the issues may be picked to study in future research.

As the computational results in this subsection suggests, a network achieved with reasonable data and a practical-size problem (see parts b and c in Fig. 6) has a configuration close to well-known ones, *e.g.*, star and triangular networks. However, the model proposed in this paper is more general than those based on pre-assigned configurations for designing RTNs. There is, indeed, no need to consider a special pre-assigned configuration when establishing a new RTN. Moreover, as already mentioned, the model introduced in this paper can be employed to analyze the effect of different changes in existing RTNs in order to improve or extend them.

6. CONCLUSION

In this study, a novel bi-objective mathematical programming model is presented for designing a rapid transit network based on a hub and spoke model. The network consists of stopovers (stations) in the hub and spoke alignments. In the proposed hub location model, both hub-level and spoke-level sub-networks are composed of multiple lines. The model relaxes some of the usual assumptions in classical hub location models by considering the possibility of the transshipment of the flows in spoke nodes (in addition to hub nodes), the setup costs for hub and spoke nodes and edges, and the capacity constraints for hub and spoke edges. In addition to determining the network infrastructure, which is done by making decisions about the location of hub and spoke nodes and the selection of hub and spoke edges, the model simultaneously determines hub and spoke lines and routes of flow to transport the demand between OD pairs. The model incorporates profit maximization (by allowing a portion of demands to be unserved) and service time minimization objectives in order to manage the interests of both the service provider and the operator.

Relaxing the assumptions and making different decisions to design a hub-based RTN is supposed to lead to better solutions for real-word situations. However, this implies a significant level of difficulty in the problem formulation as well as the computational procedure needed to solve it. Therefore, to solve realistic-size instances of the problem, a meta-heuristic algorithm was developed on the basis of an ALNS method. In the proposed solution algorithm, for an upper-level sub-problem, the network design and the line configuration decisions were managed by the ALNS with the aim of maximizing the weighted objective function. Also, in each iteration of

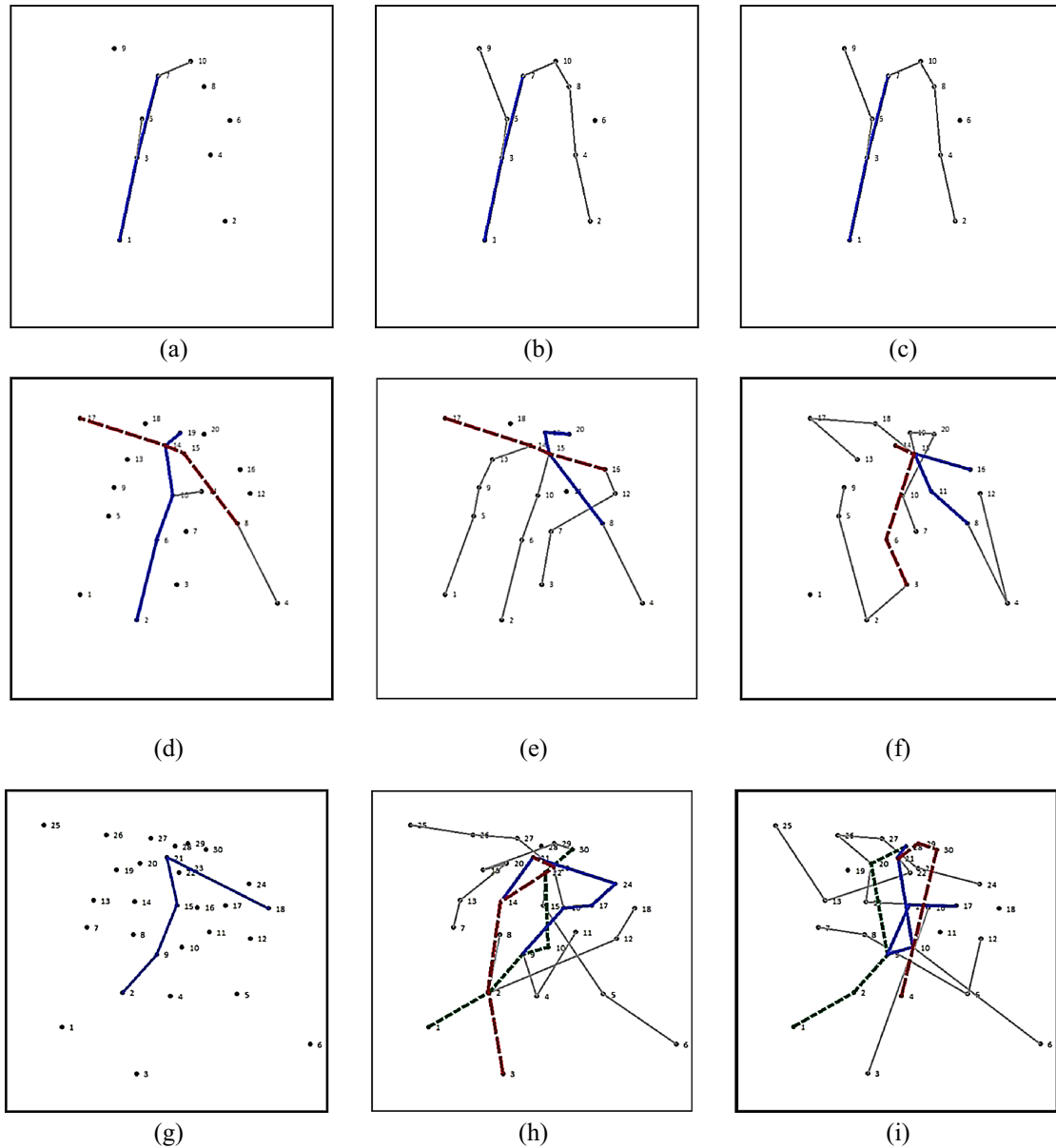


FIGURE 8. Configuration of the networks obtained for the experiments corresponding to the (a) 1st, (b) 2nd, (c) 3rd, (d) 4th, (e) 5th, (f) 6th, (g) 7th, (h) 8th, and (h) 9th rows of Table 6.

the ALNS, a lower-level sub-problem was solved by the CPLEX to make flow routing decisions for the current network solution.

In this study, some computational experiments were performed on the well-known AP and Turkish network datasets to assess the performance of the proposed model and the solution algorithm and to analyze the resulting hub networks. The performance of the ALNS algorithm was compared to that of the CPLEX solver. The computational results proved the efficiency of the proposed ASLN algorithm; it was able to find high-quality solutions within short CPU times. According to the results, the increasing difficulty of solving

larger-size instances precludes the use of commercial solvers. As another task, the resulting hub networks were analyzed under various parameter settings. The revenues and the discount factor were varied, and analyses were performed on the obtained solutions mainly in terms of the obtained values of the objective functions, the percentage of the served demands and the topology of the network. The results suggest the effectiveness and efficiency of the proposed model, and useful insights are provided into the interactions among different aspects of the studied complex decision problem.

For future research, other algorithms, either exact or meta-heuristic, can be developed, and more detailed numerical investigations are recommended so as to obtain deeper managerial insights. From the modeling point of view, even more applicable models can be found, for example, by incorporating realistic constraints (*e.g.*, the capacity limitation of stations and the constraints on the angles between the consecutive edges of transit lines), tactical decisions (*e.g.*, determining the frequencies of transit lines), elastic demands, alternative modes of transportation and factors of uncertainty.

REFERENCES

- [1] R.K. Ahuja, Ö. Ergun, J.B. Orlin and A.P. Punnen, A survey of very large-scale neighborhood search techniques. *Discrete Appl. Math.* **123** (2002) 75–102.
- [2] A. Alibeyg, I. Contreras and E. Fernández, Hub network design problems with profits. *Transp. Res. E Logist. Transp. Rev.* **96** (2016) 40–59.
- [3] T. Aykin, Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem. *Eur. J. Oper. Res.* **79** (1994) 501–523.
- [4] M. Basirati, M.R. Akbari Jokar and E. Hassannayebi, Bi-objective optimization approaches to many-to-many hub location routing with distance balancing and hard time window. *Neural. Comput. Appl.* **32** (2020) 13267–13288.
- [5] G. Bruno, G. Ghiani and G. Improta, A multi-modal approach to the location of a rapid transit line. *Eur. J. Oper. Res.* **104** (1998) 321–332.
- [6] L. Cadarso and Á. Marín, Improved rapid transit network design model: considering transfer effects. *Ann. Oper. Res.* **258** (2017) 547–567.
- [7] J.F. Campbell, A.T. Ernst and M. Krishnamoorthy, Hub arc location problems: Part I – Introduction and results. *Manage. Sci.* **51** (2005) 1540–1555.
- [8] J.F. Campbell, A.T. Ernst and M. Krishnamoorthy, Hub arc location problems: Part II – Formulations and optimal algorithms. *Manage. Sci.* **51** (2005) 1556–1571.
- [9] D. Canca, A. De-Los-Santos, G. Laporte and J.A. Mesa, An adaptive neighborhood search metaheuristic for the integrated railway rapid transit network design and line planning problem. *Comput. Oper. Res.* **78** (2017) 1–14.
- [10] D. Canca, A. De-Los-Santos, G. Laporte and J.A. Mesa, Integrated railway rapid transit network design and line planning problem with maximum profit. *Transp. Res. E Logist. Transp. Rev.* **127** (2019) 1–30.
- [11] S. Çetiner, C. Sepil and H. Süral, Hubbing and routing in postal delivery systems. *Ann. Oper. Res.* **181** (2010) 109–124.
- [12] K. Chen, X. Xin, T. Zhang and Z. Yang, Multiport cooperative location model with a safe-corridors setting in West Africa. *Int. J. Logist. Res. Appl.* **23** (2020) 580–601.
- [13] I. Contreras and M. O’Kelly, Hub location problems. In: *Location Science*, edited by G. Laporte, S. Nickel and F. Saldanha da Gama. Springer International Publishing, Cham (2019) 327–363.
- [14] I. Contreras, E. Fernández and A. Marín, The tree of hubs location problem. *Eur. J. Oper. Res.* **202** (2010) 390–400.
- [15] I. Contreras, M. Tanash and N. Vidyarthi, The cycle hub location problem. Montreal, Technical Report CIRRELT (2013).
- [16] K. Danach, S. Gelareh and R. Neamatian Monemi, The capacitated single-allocation p-hub location routing problem: a Lagrangian relaxation and a hyper-heuristic approach. *EURO J. Transp. Logist.* **8** (2019) 597–631.
- [17] A. De-Los-Santos, G. Laporte, J.A. Mesa and F. Perea, Simultaneous frequency and capacity setting for rapid transit systems with a competing mode and capacity constraints. In: *14th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems*. Vol. 42. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2014) 107–121.
- [18] A.T. Ernst and M. Krishnamoorthy, Efficient algorithms for the uncapacitated single allocation p-hub median problem. *Location Sci.* **4** (1996) 139–154.
- [19] L.F. Escudero and S. Muñoz, A survey-based approach for selecting the stations and links for a rapid transit network. *Int. J. Comput. Intell. Syst.* **7** (2014) 565–581.
- [20] R. Garcia, A. Garzon-Astolfi, A. Marn, J.A. Mesa and F.A. Ortega, Analysis of the parameters of transfers in rapid transit network design. In: *5th Workshop on Algorithmic Methods and Models for Optimization of Railways (ATMOS’05)*. Vol. 2. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2006).
- [21] S. Gelareh and S. Nickel, Hub location problems in transportation networks. *Transp. Res. E Logist. Transp. Rev.* **47** (2011) 1092–1111.
- [22] G. Gutiérrez-Jarpa, C. Obreque, G. Laporte and V. Marianov, Rapid transit network design for optimal cost and origin-destination demand capture. *Comput. Oper. Res.* **40** (2013) 3000–3009.

- [23] G. Gutiérrez-Jarpa, G. Laporte, V. Marianov and L. Moccia, Multi-objective rapid transit network design with modal competition: the case of Concepción, Chile. *Comput. Oper. Res.* **78** (2017) 27–43.
- [24] D. Huang, Z. Liu, X. Fu and P.T. Blythe, Multimodal transit network design in a hub-and-spoke network framework. *Transp. A: Transp. Sci.* **14** (2018) 706–735.
- [25] H. Karimi, The capacitated hub covering location-routing problem for simultaneous pickup and delivery systems. *Comput. Ind. Eng.* **116** (2018) 47–58.
- [26] Z. Kartal, S. Hasgul and A.T. Ernst, Single allocation p-hub median location and routing problem with simultaneous pick-up and delivery. *Transp. Res. E Logist. Transp. Rev.* **108** (2017) 141–159.
- [27] F. Kaveh, R. Tavakkoli-Moghaddam, C. Triki, Y. Rahimi and A. Jamili, A new bi-objective model of the urban public transportation hub network design under uncertainty. *Ann. Oper. Res.* **296** (2019) 131–162.
- [28] O. Kemmar, K. Bouamrane and S. Gelareh, Hub location problem in round-trip service applications. *RAIRO: Oper. Res.* **55** (2021) S2831–S2858.
- [29] S. Khosravi and M. Jokar, Many to many hub and spoke location routing problem based on the gravity rule. *Uncertain Supply Chain Manag.* **6** (2018) 393–406.
- [30] J.G. Klincewicz, Hub location in backbone/tributary network design: a review. *Location Sci.* **6** (1998) 307–335.
- [31] M. Labbé and H. Yaman, Solving the hub location problem in a star–star network. *Networks* **51** (2008) 19–33.
- [32] G. Laporte, J.A. Mesa, F.A. Ortega and I. Sevillano, Maximizing trip coverage in the location of a single rapid transit alignment. *Ann. Oper. Res.* **136** (2005) 49–63.
- [33] G. Laporte, Á. Marín, J.A. Mesa and F.A. Ortega, An integrated methodology for the rapid transit network design problem. In: *Algorithmic Methods for Railway Optimization*. Vol. 4359. Springer, Berlin (2007) 187–199.
- [34] G. Laporte, A. Marín, J.A. Mesa and F. Perea, Designing robust rapid transit networks with alternative routes. *J. Adv. Transp.* **45** (2011) 54–65.
- [35] F. López-Ramos, E. Codina, Á. Marín and A. Guarnaschelli, Integrated approach to network design and frequency setting problem in railway rapid transit systems. *Comput. Oper. Res.* **80** (2017) 128–146.
- [36] A. Mahéo, P. Kilby and P.V. Hentenryck, Benders decomposition for the design of a hub and shuttle public transit system. *Transp. Sci.* **53** (2019) 77–88.
- [37] A.İ. Mahmutogullari and B.Y. Kara, Hub location problem with allowed routing between nonhub nodes. *Geog. Anal.* **47** (2015) 410–430.
- [38] Á. Marín, An extension to rapid transit network design problem. *TOP* **15** (2007) 231–241.
- [39] Á. Marín and R. García-Ródenas, Location of infrastructure in urban railway networks. *Comput. Oper. Res.* **36** (2009) 1461–1477.
- [40] E. Martins de Sá, I. Contreras and J.-F. Cordeau, Exact and heuristic algorithms for the design of hub networks with multiple lines. *Eur. J. Oper. Res.* **246** (2015) 186–198.
- [41] E. Martins de Sá, I. Contreras, J.-F. Cordeau, R.S. de Camargo and G. de Miranda, The hub line location problem. *Transp. Sci.* **49** (2015) 500–518.
- [42] G. Nagy and S. Salhi, The many-to-many location-routing problem. *TOP* **6** (1998) 261–275.
- [43] S. Nickel, A. Schöbel and T. Sonneborn, Hub location problems in urban traffic networks. In: *Mathematical Methods on Optimization in Transportation Systems*, edited by M. Pursula and J. Niittymäki. Springer, Boston (2001) 95–107.
- [44] F.A. Oliveira, E.M. de Sá and S.R. de Souza, Benders decomposition applied to profit maximizing hub location problem with incomplete hub network. *Comput. Oper. Res.* **142** (2022) 105715.
- [45] D. Pisinger and S. Ropke, Large Neighborhood Search. In: *Handbook of Metaheuristics*, edited by M. Gendreau and J.-Y. Potvin. Springer International Publishing, Cham (2019) 99–127.
- [46] H.M. Repolho, A.P. Antunes and R. L. Church, Optimal location of railway stations: the Lisbon-Porto high-speed rail line. *Transp. Sci.* **47** (2013) 330–343.
- [47] S. Ropke and D. Pisinger, An adaptive large neighborhood search heuristic for the pickup and delivery problem with time windows. *Transp. Sci.* **40** (2006) 455–472.
- [48] C. Shushan, L. Qinghuai and S. Zhong, Design of urban rail transit network constrained by urban road network, trips and land-use characteristics. *Sustainability* **11** (2019).
- [49] G. Taherkhani and S.A. Alumur, Profit maximizing hub location problems. *Omega* **86** (2019) 1–15.
- [50] P.Z. Tan and B.Y. Kara, A hub covering model for cargo delivery systems. *Networks* **49** (2007) 28–39.
- [51] K. Tavassoli and M. Tamannaie, Hub network design for integrated Bike-and-Ride services: a competitive approach to reducing automobile dependence. *J. Cleaner Prod.* **248** (2020) 119247.
- [52] T. Thomadsen and J. Larsen, A hub location problem with fully interconnected backbone and access networks. *Comput. Oper. Res.* **34** (2007) 2520–2531.
- [53] H. Tikani, M. Honarvar and Y.Z. Mehrjerdi, Developing an integrated hub location and revenue management model considering multi-classes of customers in the airline industry. *Comput. Appl. Math.* **37** (2018) 3334–3364.
- [54] A. Verma, A. Kumari, D. Tahlyan and A.B. Hosapujari, Development of hub and spoke model for improving operational efficiency of bus transit network of Bangalore city. *Case Stud. Transp. Policy* **5** (2017) 71–79.
- [55] S.T. Windras Mara, R. Norcahyo, P. Jodiawan, L. Lusiantoro and A.P. Rifai, A survey of adaptive large neighborhood search algorithms and applications. *Comput. Oper. Res.* **146** (2022) 105903.

- [56] J. Wu, R. Song, Y. Wang, F. Chen and S. Li, Modeling the Coordinated operation between bus rapid transit and bus. *Math. Probl. Eng.* **2015** (2015) 709389.
- [57] H. Yaman, B.Y. Kara and B.Ç. Tansel, The latest arrival hub location problem for cargo delivery systems with stopovers. *Transp. Res. B: Methodol.* **41** (2007) 906–919.

Subscribe to Open (S2O)

A fair and sustainable open access model



This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

Please help to maintain this journal in open access!

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting subscribers@edpsciences.org

More information, including a list of sponsors and a financial transparency report, available at: <https://www.edpsciences.org/en/maths-s2o-programme>