University of Tehran Press

Industrial Management Journal

Online ISSN: 3115 - 7386

Homepage: https://imj.ut.ac.ir

Many-to-many location-routing problem with multiple paths, heterogeneous vehicles, and time windows

Farzad Bahrami [1]

1. Corresponding author, Assistant prof., Department of Industrial Management, Faculty of Administrative Sciences and Economics, Arak University, Arak, Iran. E-mail: f-bahrami@araku.ac.ir

Article Info

Article type:
Research Article

Article history:
Received May 23, 2025
Received in revised form September 13, 2025
Accepted October 02, 2025
Available online October 25, 2025

Keywords:
Logistics; hub location-routing; parcel delivery services; meta-heuristics.

ABSTRACT

Objective: This study introduces a location-routing model tailored for parcel delivery in large, sparsely populated regions with limited infrastructure. It aims to minimize system costs by optimizing hub placement, city-to-hub assignments, routing paths, and fleet composition. The model accounts for real-world complexities such as diverse vehicle types, flexible delivery time windows, and multiple pickup/delivery paths, offering a strategic planning tool for logistics operations in challenging environments.

Methodology: To solve this NP-hard problem, the researchers reformulated a mixed-integer nonlinear program (MINLP) into a more computationally efficient mixed-integer programming (MIP) model. For larger instances, they developed a two-stage hybrid metaheuristic: the first stage uses an Artificial Bee Colony (ABC) algorithm to explore hub locations and initial allocations, while the second stage applies Simulated Annealing (SA) with local search to optimize routing and assignments. Validation was performed using CPLEX for small instances and benchmarked against a published SA-based method across 75 test scenarios and two real-world case studies from an Iranian parcel delivery company.

Results: The hybrid method achieved optimal or near-optimal solutions faster than CPLEX for minor problems and outperformed the SA benchmark for larger ones, improving solution quality by 4% and reducing routes by 11%. The model also increased 24-hour deliveries by 4% without raising costs. The SA phase alone contributed a 1.6% cost reduction by restructuring the network. Case studies confirmed the model's practical value, consistently identifying robust hub configurations across diverse network scales and operational strategies.

Conclusion: This study presents a strategic planning tool for parcel delivery in challenging geographic and infrastructural conditions. It enables logistics managers to minimize operational costs while maintaining stable hub configurations during network expansion. A case study in Iran highlights its long-term value: a four-hub network with a 680 km line-haul limit offers superior nationwide coverage compared to a three-hub setup with a 510 km limit focused on major cities.

Cite this article: Bahrami, F., (2025). Many-to-many location-routing problem with multiple paths, heterogeneous vehicles, and time windows. *Industrial Management Journal*, 17(4), 100-139. https://doi.org/10.22059/imj.2025.402989.1008266

Publisher: University of Tehran Press.

DOI: https://doi.org/10.22059/imj.2025.402989.1008266

Introduction

Distribution network design is a critical logistics challenge significantly affecting service continuity and customer satisfaction. Logistic service providers in parcel services confront a more complex situation than other services, as they serve many small-sized customers, each of whom may send or receive parcels. The delivery network design's most important part is locating hubs and determining distribution routes between nodes and hubs (Sharma et al., 2025). Since 2011, controlling pollution sources and incorporating environmental considerations have also drawn researchers' attention (Nielsen et al., 2024). Hubs are facilities that merge, connect, and exchange flow points between demand origins and destinations (O'Kelly et al., 2025). They are utilized to group parcel volumes at demand nodes, enabling cost savings through economies of scale in less-than-truckload (LTL) transportation (Lüer-Villagra et al., 2019).

This paper presents a new model for designing a network for an investor who decided to run a new door-to-door parcel delivery system in Iran's broad and sparsely populated country (the 18th largest in the world). Iran is a mountainous country, whose western landscape is dominated by uneven mountain ranges that separate the west from the east. (Bahrami et al., 2016). The northern part of Iran is covered by dense rainforests with populated and close cities, and the eastern part consists mainly of desert basins with sparse and far cities. The widespread practice of parcel delivery in Iran is 24, 48, and 72 hours because the distance between the farthest cities is more than 2600 km. Incomplete hub location problems have become a significant focus in realistic network design. Starting from early models, such as the incomplete hub covering (Athamneh et al., 2023) to more advanced formulations like the uncapacitated single-allocation p-hub median (Brimberg et al., 2021; De Camargo et al., 2017), researchers have employed heuristic methods (Andaryan et al., 2024; Davari et al., 2013), decomposition techniques (de Sá, Morabito, & de Camargo, 2018), and uncertainty modeling to address practical constraints. In a similar vein, (Öztürk et al., 2021) Presented a refined methodology for solving large-scale incomplete p-hub median problems and highlighted future directions for extending the model to other hub-location variants.

However, a research gap exists to consider the real operating conditions of various vehicle types and time windows in a network with incomplete connections and sparse cities, such as Iran's road network. In short, the upfront problem aims to answer the following questions:

- If all hubs should be linked to each other and transfer their parcels in a determined time window, what is the optimal number of hubs and where should these hubs be opened?
- What is the optimal set of cities assigned to each hub, and which should be linked to each hub in a single assignment network?

* How many routes are needed for each hub, and what is the best way to assign nodes to routes?
* How many time periods should be considered in routing parts, and how can the proposed system reduce the delivery time?
* If heterogeneous vehicles exist, how many vehicles of each type are required?

The proposed network design problem can be classified in two separate ways: (1) a Latest Arrival HLP (LAHLP) with multiple stopovers (Yaman et al., 2007), or (2) a many-to-many location-routing problem (MMLRP) (Nagy & Salhi, 1998). In the LAHLP with stopovers, researchers incorporate waiting times at hubs and consider ground transportation with multiple stopovers in non-hub nodes while all hub nodes are directly connected. While in an MMLRP, several clients may send products to each other, as all of them potentially have pickup and delivery, to find the location of hubs and transportation flow among the clients (i.e., non-hub nodes).

Although the assumptions of LAHLP with stopovers are considered, this paper adopts the MMLRP format. This choice enables (1) to use routing features in the location problem (except tour constraints), (2) to define heterogeneous vehicles and time windows, and (3) to model the location-routing problem (LRP) differently. As noted in (Gu et al., 2024), the many-to-many multi-commodity routing problem is sufficiently broad to encompass split pickups, multi-stage deliveries, and unified vehicle routing optimization. Furthermore, a partially connected network can compel designers to adopt multi-path routing in the MMLRP instead of simple tours.

To fulfill the promises of the common delivery time, the service provider sets time windows for deliveries and pickups in the routing and line haul transport (i.e., hub-to-hub connections). Each hub may need to manage more than one route, and each route may use various vehicles with different capacities and travel expenses. Such a fundamental assumption of Iran turns a well-known MMLRP into a many-to-many location multi-path routing problem with heterogeneous vehicles and time windows.

The MMLRP was shown to be NP-hard by Wasner and Zäpfel (2004) as it combines the facility location and the traveling salesman problem. However, the multi-path routing of this paper has a more complex solution space, as each hub can handle more than one route with different lengths and nodes. Solving such problems by exact algorithms is computationally intolerable, so for medium and large instances, the model is solved by a two-stage hybrid algorithm based on different meta-heuristics. Since the problem is relatively new, 75 test instances are generated based on the practical assumptions, with a node size ranging from 10 to 50. Moreover, three area dimensions are considered to mimic the situation of a broad and sparse country with different breadth and number of cities, and to indicate the applicability of the presented model and proposed solution method. Similar to the Bahrami et al. (2016), a real case with all 31 capital cities of Iran provinces is considered. As the company may expand its coverage, the proposed solution method also considers and solves the second case with 59 cities.

The remainder of the paper is organized as follows. In Section 2, the literature is analyzed in detail. Section 3 describes the problem and a mixed-integer non-linear programming (MINLP) model. Also, a lemma is proposed to linearize the relations of heterogeneous vehicles. Due to the NP-hard nature of this problem, a new two-stage hybrid method based on two meta-heuristics is designed in Section 4. In Section 5, experimental analyses on the generated test instances and two real-world cases show the efficiency of the proposed methodology. Finally, the summary and conclusions are discussed in the last section.

Literature Background

The Hub Location Problem (HLP) is a foundational area in location theory and logistics that centers on optimizing hub-and-spoke networks. At its core, HLP seeks to determine the optimal placement of hubs and the spokes assignment, balancing strategic (hub selection) and tactical (spoke allocation) decisions. Traditional HLP often assumes static connections between spokes and their assigned hubs. However, this can limit real-world logistics, particularly in scenarios such as LTL services where routes are not fixed and circular pickup and delivery tours are necessary. These practical realities have driven the need for more flexible and dynamic modeling frameworks, as pointed out by O’Kelly and colleagues (O'Kelly, 1986; O'Kelly et al., 2015). For instance, Estrada-Romeu and Robusté (2015) considered the HLP with stopovers to devise a methodology for identifying cost-efficient merging strategies in LTL systems, finding that their approach could reduce transportation costs by up to 20%. As hub selection and spoke allocation underpin the efficiency of wide-ranging networks—including postal, transport, and telecom systems—HLP continues to play a critical role in reducing operational complexities and costs (Wandelt et al., 2025).

Building on HLP, Hub Location Routing Problems (HLRP) have emerged to jointly optimize hub locations and the routing of vehicles that serve customers. Unlike classical HLP, HLRPs allow for greater practical realism by permitting multi-stop, circular, or dynamic vehicle routes instead of fixed spoke-to-hub assignments. This flexibility is essential for capturing real-world distribution and freight system operations. However, expanding the problem in this way introduces new challenges: HLRPs must simultaneously account for vehicle and hub capacity constraints, multiple allocation strategies for non-hub nodes, and the evolution of network parameters over time—a feature largely absent from static models (Ratli et al., 2022; Wu et al., 2022). While incremental research has begun to address these areas, most existing models still lack full integration of simultaneous hub and vehicle capacity constraints, multi-assignment of nodes, and multi-period dynamics (Aloullal et al., 2023) .

Taking integration even further, the Multi-Modal Location Routing Problem (MM-LRP) combines hub location with the vehicle routing problem in multi-modal networks. In MM-LRP, multiple hubs must be located, and routes must be designed so customers can serve as shippers and receivers—in essence, enabling direct flows between any pair of customers. This model minimizes total system costs, including hub installation, transportation, and routing. It also ensures compliance with operational and design constraints (Abbasi et al., 2019). Unlike classical models, MM-LRP tightly couples the design of hierarchical networks with real-world vehicle routing, improving both cost efficiency and service levels (Gianpaolo et al., 2013; Shi et al., 2023). This holistic approach better mirrors the complex supply chains in practice and sets the stage for further technical advancements.

A significant advance in location-routing research has been modeling heterogeneous vehicle fleets, which introduces substantial operational realism and complexity. Early studies relied on homogeneous vehicles, but more recent works (e.g., Wang & Li, 2017; Zhao et al., 2018) have explored the impact of joint delivery alliances and low-carbon logistics solutions in the context of diverse vehicle types. These models integrate simultaneous pickup and delivery operations and time window constraints to manage emissions and costs effectively.

Time window constraints in vehicle routing problems have evolved from strict hard windows requiring service within fixed intervals (Basirati et al., 2020; Kartal et al., 2017) to more flexible multi-period and dynamic scheduling approaches (Aloullal et al., 2023). Additionally, queueing theory has been utilized to model waiting times under uncertainty (Pourmohammadi et al., 2023). However, standardized classifications and comprehensive frameworks for different types of time window constraints remain insufficiently developed (Fallah-Tafti et al., 2022; Ghodratnama et al., 2013).

Significant methodological advances have been made in solving location-routing and hub location problems in recent years. Advanced exact optimization methods, such as branch-price-and-cut, now provide precise solutions that can accommodate complex features like heterogeneous fleets and environmental constraints. This represents a natural progression in addressing the complex challenges of modern urban logistics (Wang et al., 2025).

On the other hand, tackling the considerable scale and uncertainty of real-world problems often necessitates heuristic and metaheuristic methods. Genetic algorithms (Wang et al., 2023) , simulated annealing (Oudani et al., 2023) , Tabu search (Bütün et al., 2021), and hybrid matheuristics (Aloullal et al., 2023) are commonly employed to search vast solution spaces efficiently. For example, in the context of the MMLRP, Rieck, Ehrenberg and Zimmermann (2014) considered a variant for the timber-trade industry, proposing an MIP model for small-scale instances and a genetic algorithm (GA) to solve large-scale ones. Similarly, simulated annealing (SA) has proven effective for MMLRPs. Karaoglan et al. (2012) developed a two-phase SA-based heuristic for the LRP with simultaneous pickup and delivery (LRPSPD), while Bahrami et al. (2016) presented a multi-step SA method to maximize profit for door-to-door delivery services in Iran. Bahrami et al. (2017) addressed a hub location-routing problem for sparse networks in Iran by designing a two-stage method; it first uses a genetic algorithm to locate hubs and allocate nodes, then employs simulated annealing to optimize the resulting vehicle routes.

Other recent applications include fuzzy multi-objective models that address flow and cost uncertainties (Pourmohammadi et al., 2023), while stochastic techniques improve ambulance routing under uncertainty (Khoshgehbari & Al-e, 2023). Other recent applications include carbon emission minimization in green logistics (Kabadurmus & Erdogan, 2023), Q-learning-based heuristics for sustainable waste collection (Shang et al., 2023), hybrid metaheuristics for urban microhub optimization (Guo et al., 2024), and attention-enhanced simulated annealing for last-mile electric vehicle delivery (Zhao et al., 2025). These diverse approaches address contemporary logistics networks' dynamic and multifaceted nature. A summary of selected studies is presented in Table 1 for an overview of key research.

While location-routing models have advanced, a gap remains in integrating multiple complex, real-world constraints. This study addresses this by introducing a new MMLRP variant. Tailored for parcel networks in large countries with sparse or incomplete road infrastructure, the model uniquely integrates a heterogeneous vehicle fleet, service time windows, and multiple routing paths. To solve this NP-hard problem, a mixed-integer programming model is formulated, and a two-stage hybrid metaheuristic (Artificial Bee Colony and SA) is proposed for large-scale, real-world instances. This integrated approach offers a more realistic framework for modern logistical challenges.

Table 1. A review of key studies conducted in the field of Hub Location-Routing Problems.

Study	Problem Class	Fleet	Capacity	Time Windows	Pickup and Delivery	Uncertainty	Solution Method
Karaoglan (2012)	LRPSPD	Ho	V/De	-	SPD	Det	Heu
Ghodratnama (2013)	Hub Covering	He	V/Fa	-	-	Fu	E
O'Kelly, Campbell (2015)	HLRP (Network Design)	-	-	-	-	Det	E
Rieck (2014)	Many-to-Many LRP	Ho	V	-	Separate	Det	Heu
Estrada-Romeu (2015)	LTL Long-Haul Routing	-	-	-	-	Det	Heu
Bahrami (2016)	Integrated HLRP	Ho	V	-	-	Det	Heu
Bahrami (2017)	Ground HLRP (LTL)	Ho	V	-	Separate	Det	Heu
Kartal (2017)	pHLRP-SPD	Ho	-	-	SPD	Det	Heu
Wang and Li (2017).	LCLRP	He	V	Ha	SPD	Det	Heu
Zhao, Wang (2018).	2E-CLRP	He	V/ID	-	Delivery Only	Det	Heu
Basirati (2020)	MMHLRP	Ho	V/Hu	Ha	SPD	Det	E - Heu
Abbasi (2019)	Large-Scale MMHLRP	Ho	-	TT	-	Det	Heu
Bütün (2021)	DCHC (HLRP)	Ho	Hu	-	-	Det	Heu
Pourmohammadi (2023)	HLRP	-	-	-	-	Fu	Heu
Fallah-Tafti (2022)	CLSC Network Design	-	Fa	-	P&D	Fu	Heu
Ratli, Urošević (2022)	pHLRP	Ho	-	TN	-	Det	Heu
Wu (2022)	MAHLRP	Ho	V/Hu	-	SPD	Det	Heu
Aloullal (2023)	Multi-period HLRP	-	-	-	-	Det	Heu
Kabadurmus (2023)	Green VRP (GVRP)	He	V	-	Split Delivery	Det	Heu
Khoshgehbari (2023)	ALRP	He	FS	RT	-	St	E - Heu
Shang (2023)	Green LRP (GLRP)	He	V	-	-	Det	Heu
Shi (2023)	LRP (Recycling)	Ho	V/Fa	-	SPD	Det	Heu
Wang (2023)	Robust HLRP	-	-	-	-	Ro	E
Guo (2024)	MDLRP with Microhubs	-	V	-	Mixed P&D	Ro	Heu
Wang (2025)	LRP	He	V/De	-	-	Det	E
Zhao (2025)	LRP-GLD	Ho	V-Ba	-	Delivery Only	Det	Heu
This paper	MMHLRP (Multi-path)	He	V	Ha	SPD	Det	Heu

* **2E-CLRP**: Two-Echelon Capacitated LRP; **ALRP**: Ambulance LRP; **Ba**: Battery; **CLSC Network Design**: Closed-Loop Supply Chain; **DCHC (HLRP)**: Directed Cycle Hub location and routing problem under Congestion; **De**: Depot; **Det**: Deterministic; **E**: Exact; **Fa**: Facility; **FS**: Fleet Size; **Fu**: Fuzzy; **Green VRP (GVRP)**: Green Vehicle Routing Problem; **H**: Hard; **Ha**: Hard; **He**: Heterogeneous; **Heu**: Heuristic; **Ho**: Homogeneous; **Hu**: Hub; **ID**: Intermediate Depots; **LCLRP**: Low-Carbon LRP; **LRP**: Location-Routing Problem; **LRP-GLD**: LRP for Green Last-mile Delivery; **LRPSPD**: LRP with Simultaneous Pickup and Delivery; **LTL**: Less-than-Truckload; **MAHLRP**: Multi-Allocation HLRP; **MDLRP with Microhubs**: Multi-Depot LRP; **MMHLRP**: Many-to-Many HLRP; **P&D**: Pickup and Delivery; **pHLRP**: p-Hub median LRP; **pHLRP-SPD**: p-Hub median LRP with Simultaneous Pickup and Delivery; **Ro**: Robust; **RT**: Response Time; **SPD**: Simultaneous Pickup and Delivery; **St**: Stochastic; **TN**: Tour Nodes; **TT**: Tour Time; **V**: Vehicle.

Materials and Methods

Problem definition and formulation

This section first describes the problem and introduces the system assumptions in Iranian parcel delivery services. Afterwards, the MMLRP with multi-path routing, distinct types of vehicles, and time windows is modeled as a MINLP formulation. Finally, the heterogeneous vehicles' conditions are linearized by four constraints to form the polynomial-sized MIP formulation.

Problem description and assumptions

The problem can be defined as follows. Let G= (N, A) be a network, where N is a set of nodes and A= (i, j): i, j∈N is the set of arcs that is in a non-directed form. For each (i, j), a route distance/time, $T_{ij} = T_{ji}$ is considered, and the total of D_{ij} parcels should be transferred from node i to node j. To bundle the parcels and take advantage of economies of scale, at least one hub should be selected from N hubs. The capacity of hubs is unlimited; however, a fixed cost (CH_i) is associated with each hub node, in which all nodes are potential hubs. Besides, each node should be connected to one and only one hub (i.e., single assignment); however, hubs are fully connected.

In the proposed problem, each hub must open routes for collecting and distributing its assigned nodes. In fact, $k \in N_k$ ($N_k \leq N-1$) routes are established to collect and deliver parcels between nodes and hubs. When k=1, the problem converts to the vehicle routing problem (VRP), and with k=N-1, the problem becomes the 1-hub location. A fixed cost (R) is related to establishing each route.

In our special real application, each route starts from and ends at a hub, but in a path (not necessarily a tour). Besides, more than one route can start from a hub (i.e., multi-path routing); however, pickup and delivery of each route are not simultaneous. In fact, deliveries happen first, and pickups occur when vehicles return to their hubs. Each route consists of $p \in N_p$ ($N_p \leq N$) places so that hubs can only take the first position of one or more routes; however, non-hub nodes should take only one place in one route except for the first place.

In the proposed problem, transportation costs are not calculated based on the traveling distance, as it assumes vehicles are in service for the entire day. Also, $v \in N_v$ different types of road vehicles with different capacities carry out all system transportation. The number of vehicles has no limit, as the company can rent as many vehicles as it needs; however, variable cost (CV_v) and capacity (Q_v) are considered for the v-th kind of vehicles. In this system, it is economical to ask drivers to stay at the last nodes of the route for some hours or days and collect the parcels of the visited nodes. The number of each kind of vehicle in each route depends on the maximum number of bundled parcels picked up or delivered on the route. This assumption is irrelevant for line-haul connections, and the number of vehicles between two hubs can differ in one direction.

Routes are considered to be in two periods: regular and extended. In expected conditions, vehicles can travel a maximum distance/time of TR1 in a route, which provides less than 24 hours of traveling between two nodes. In the extended time, the vehicles use the extra amount of distance/time, R2 , which causes the traveling time between two nodes to exceed 24 hours. The total collected and delivered parcels of a route that violate the expected time include a penalty cost (CP) to provide additional finance for vehicles traveling through the extended time routes (or extended routes). We assume only one time period (TH) for line haul transportation since all hubs are connected and parcel transfer is fast enough. Therefore, the company has three time windows of 24, 48, and 72 hours for the delivery of parcels between two cities, based on two time periods for routes and a time period for line-hauls. This assumption will be explained later in more detail.

In delivery services, a market usually determines the rate of services between cities, so the logistic service provider prefers to minimize its total fixed and variable costs. To satisfy this goal, managers should make interrelated decisions about (1) the number and location of hubs, (2) allocation of nodes (cities) to hubs, (3) transportation routes connecting nodes to hubs, and (4) the number of different vehicles in routes and line hauls.

Problem formulation

Expanding the models presented by Wasner and Zäpfel (2004), Karaoglan et al. (2012), Bahrami et al. (2016)We propose a new MMLRP with multi-path routing, time windows, and heterogeneous vehicles. While the concept of heterogeneous vehicles is mainly taken from (Čupić & Teodorović, 2014)Its formulation is optimally linearized for different vehicle types. Table2**Error! Reference source not found.** presents the variables of the proposed model.

Table2. Variables

	Notations	Description
	Decision Variables	
	x_{pik}	1, if node i is allocated to the p-th place of the k-th route; and 0, otherwise.
	y_{ij}	1 if node i is allocated to hub j; and 0, otherwise.
	B_k	1, if the distance/time of route k is more than the normal time; and 0, otherwise.
	VR_{vk}	Number of the v-th kind of vehicles in route k
	VH_{vij}	Number of the v-th kind of vehicles between nodes i and j
	Additional Variables	
	MRP_k	Amount of the total collected parcels in the k-th route
	MRD_k	Amount of the total delivered parcels in the k-th route
	MMR_k	Maximum number of parcels in the traffic of the k-th route
	MHH_{ij}	Amount of the transferred parcels from node i to j

In the proposed problem, more than one route may be assigned to a hub, unlimited heterogeneous vehicles can be selected for routes and line hauls, and all parcels of the routes and hub nodes should be delivered by vehicles of the line hauls. Therefore, it is useless for the problem to be modeled based on routes as its variables need more than six indices in a linearized framework. So, two variables with three and two indices are used to model the problem. Model I is as follows:

Model I:

$$\begin{aligned}\text{Min } Z = & \sum_i y_{ii} \,.\, CH_i + \sum_i \sum_k x_{1ik} \,.\, CR + \sum_v \sum_k VR_{vk} \,.\, CV_v + \\ & \alpha \sum_v \sum_i \sum_j VH_{vij} \,.\, CV_v + \sum_k B_k \,.\, (MRP_k + MRD_k) \,.\, CP\end{aligned} \tag{1}$$

s.t.

$$\sum_j y_{ij} = 1 \qquad \forall i \in N \tag{2}$$

$$y_{jj} - y_{ij} \geq 0 \qquad \forall i, j \in N \tag{3}$$

$$\sum_i x_{pik} \leq 1 \qquad \forall p \in N_p, \forall k \in N_k \tag{4}$$

$$\sum_i x_{pik} - \sum_i x_{p+1,ik} \geq 0 \qquad \forall p \in N_p, \forall k \in N_k \tag{5}$$

$$\sum_{p>1} \sum_k x_{pik} \leq 1 \qquad \forall i \in N \tag{6}$$

$$\sum_p \sum_k x_{pik} \geq 1 \qquad \forall i \in N \tag{7}$$

$$x_{1ik} - y_{ii} = 0 \qquad \forall i \in N, \forall k \in N_k \tag{8}$$

$$\sum_{p>1} x_{pik} - \sum_{j \neq i} y_{ij} \leq 0 \qquad \forall i \in N, \forall k \in N_k \tag{9}$$

$$\sum_{p>1} x_{pik} + x_{1ik} - y_{ij} \leq 1 \qquad \forall i, j \in N, (i \neq j), \forall k \in N_k \tag{10}$$

$$\sum_p \sum_i \sum_j x_{pik} \,.\, x_{p+1,jk} . T_{ij} \leq TR1 + B_k . TR2 \qquad \forall k \in N_k \tag{11}$$

$$y_{ii} . y_{jj} . T_{ij} \leq TH \qquad \forall i, j \in N \tag{12}$$

$$MRP_k = \sum_{p>1} \sum_i (x_{pik} \,.\, \sum_j D_{ij}) \qquad \forall k \in N_k \tag{13}$$

$$MRD_k = \sum_{p>1}\sum_i (x_{pik} . \sum_j D_{ji}) \qquad \forall k \in N_k \tag{14}$$

$$MMR_k = \max\{MRP_k, MRD_k\} \qquad \forall k \in N_k \tag{15}$$

$$MMH_{ij} = \sum_l \sum_d y_{li} . y_{dj} . D_{ld} \qquad \forall i, j \in N \tag{16}$$

$$VR_{vk} = \begin{cases} 1 \text{ if } Q_{v-1} \leq MMR_k \leq Q_v \text{ or } \left(MMR_k > Q_q \text{ and } Q_{v-1} < MMR_k - Q_q \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor \leq Q_v\right) \\ 0 \qquad \text{otherwise} \end{cases} \quad \forall v \in N_v, v \neq q, \forall k \in N \tag{17}$$

$$VR_{qk} = \begin{cases} 0 & \text{if } MMR_k \leq Q_{q-1} \\ \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor & \text{if } MMR_k > Q_{q-1} \text{ and } MMR_k - Q_q \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor \leq Q_{q-1} \\ \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor + 1 & \text{if } MMR_k > Q_{q-1} \text{ and } MMR_k - Q_q \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor > Q_{q-1} \end{cases} \quad \forall k \in N_k \tag{18}$$

$$VH_{vij} = \begin{cases} 1 \text{ if } Q_{v-1} \leq MMH_{ij} \leq Q_v \text{ or } (MMH_{ij} > Q_q \text{ and } Q_{v-1} < MMH_{ij} - Q_q \left\lfloor {}^{MMH_{ij}}/_{Q_q} \right\rfloor \leq Q_v) \\ 0 \qquad \text{otherwise} \end{cases} \quad \forall v \in N_v, v \neq q, \forall\, i, j \in N \tag{19}$$

$$VH_{qij} = \begin{cases} 0 & \text{if } MMH_{ij} \leq Q_{q-1} \\ \left\lfloor {}^{MMH_{ij}}/_{Q_q} \right\rfloor & \text{if } MMH_{ij} > Q_{q-1} \text{ and } MMH_{ij} - Q_q \left\lfloor {}^{MMH_{ij}}/_{Q_q} \right\rfloor \leq Q_{q-1} \\ \left\lfloor {}^{MMH_{ij}}/_{Q_q} \right\rfloor + 1 & \text{if } MMH_{ij} > Q_{q-1} \text{ and } MMH_{ij} - Q_q \left\lfloor {}^{MMH_{ij}}/_{Q_q} \right\rfloor > Q_{q-1} \end{cases} \quad \forall i, j \in N \tag{20}$$

where $\left\lfloor {}^{A}/_{B} \right\rfloor = \max\{n \in \mathbb{Z} \mid n \leq ({}^{A}/_{B})\}$ ($\mathbb{Z}$ is the set of integers).

$$x_{pik} = \{0,1\} \qquad \forall p \in N_p, \forall i \in N, \forall k \in N_k \tag{21}$$

$$y_{ij} = \{0,1\} \qquad \forall i, j \in N \tag{22}$$

$$B_k = \{0,1\} \qquad \forall k \in N_k \tag{23}$$

$$VR_{vk} \text{ is integer} \qquad \forall v \in N_v, \forall k \in N_k \tag{24}$$

$$VH_{vij} \text{ is integer} \qquad \forall v \in N_v, \forall i, j \in N \tag{25}$$

$$MRP_k,\ MRD_k,\ MMR_k \geq 0 \qquad \forall k \in N_k \tag{26}$$

$$MMH_{ij} \geq 0 \qquad \forall i, j \in N \tag{27}$$

The parcel delivery company needs to minimize its total cost by the objective function. (1). In the proposed network, the number of hubs is not determined in advance, so the cost of opening each hub is considered in the first term of (1). The second term measures the cost of opening each route to prevent the model from opening single-node routes. Transportation is the most important part of each logistic system and is considered in the third and fourth terms. The third term is related to the route vehicles, which will be recruited all day for pickup and delivery. However, the fourth term is associated with those of line hauls and is multiplied by the parameter α (e.g., $\alpha = 0.5$), since line haul vehicles are used for a half line. The last term of the objective function forces the model to consider 24 hours of transportation. If possible, it provides additional finance for vehicles that may travel through the extended routes, as the third term considers the cost of route vehicles only for one day.

As mentioned before, there is no restriction on the number of optimal hubs. Constraints (2) and (3) ensure that each non-hub node is assigned to one node only when it was formerly selected as a hub. Constraint (4) ensures that each numerical place of each route is assigned to at most one node. Constraint (5) assigns the route places to the nodes in the numerical order, only if a node occupies a prior place in the route. The first place of each route belongs to a hub node; however, other places belong to non-hub nodes. Each hub can start as many routes as it needs; however, non-hub nodes can only occupy one numerical place on one route. Constraints (6) and (7) reflect assumptions of nodes; if node i is a non-hub node, it can only select one place of one route, except the first place of the route, and if node i is a hub node, it can select only the first place of as many routes as it requires. To establish a proper connection between the variables of hub location-allocation and path-routing, Constraints (8) - (10) should be added to the model. Constraint (8) ensures that nodes in the first place of routes are selected as hubs. Constraints (9) and (10) check that other nodes in other places of routes are allocated to the right hubs.

Proposition. In the proposed model, Constraints (2) - (10) do not allow any illegal allocation of nodes.

Proof. Two central illegal allocations can occur: (1) a node is allocated to two places in a path, two paths of a hub, or two paths of two hubs, and (2) a node is selected as a hub and a non-hub node at the same time.

Condition (1): Consider that the first place of x_{pik} belongs to hubs, and each hub can take more than one first place. Constraint (6) states that at most one place in paths can be allocated to a non-hub node. As an allocation occurs, Constraint (5) makes sure that a node is allocated to the first place as a hub node. Besides, Constraints (8), (9) and (10) guarantee a path to be allocated to only one hub. So, this condition will not happen.

Condition (2): The strength of these constraints can be verified by contradiction. Consider the illegal allocation of an arbitrary node s as a hub and a non-hub node. Since $x_{1sw} = 1$ and $x_{lsw'} = 1$ (l>1), Constraints (2), (3) and (8) imply that $y_{ss} = 1$, while $y_{sg} = 0$ ($s \neq g$). Although the supposed variables do not violate the Constraint (6) and (7) by introducing Constraint (9) for all routes, it is obtained that $\sum_{p>1} x_{psw'} = 0$, which contradicts the assumption of $x_{lsw'} = 1$ (l>1).

As mentioned before, a service provider should handle three-time windows of 24-, 48-, and 72-hour transportation. Constraints (11) and (12) show time period restrictions in routes and line-hauls, respectively. These constraints divide the delivered promises between routes and line-hauls, considering three parameters: TR1, TR2, and TH. Figure 1 gives more explanation for Constraints (11) and (12). We know that maximum time periods in line-hauls should be less than TH, and each hub can support more than one route. When the total travel time in a route is less than TR1 (i.e., white routes), the parcels in this route will be collected and delivered in the expected time and the route is called regular route; however, when time in a route is more than TR1 and less than TR2 (i.e., yellow routes), the route is called extended route. Now, suppose both origin and destination nodes belong to the regular routes (e.g., Nodes 10 and 22). In that case, the parcels will be delivered in less than 24 hours, if one route belongs to the extended routes (e.g., Nodes 3 and 28), the parcels will be delivered in less than 48 hours, and if both nodes belong to the extended routes (e.g., Nodes 4 and 35), the parcels will be delivered in less than 72 hours.

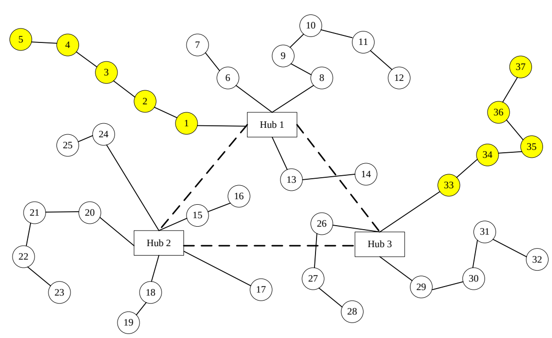

Figure 1. Network with 40 nodes and three different time windows

The number of parcels delivered and collected affects the number of each kind of vehicle on the route. Relations (13) and (14) They are related to the pickup and delivery of parcels in the routes. To balance the number of vehicles in the routes, only the maximum number of Relations (13) and (14) should be considered as a Relation (15). The Relation can easily calculate the amount of delivered parcels in line-hauls (16). Since all hubs are connected and line haul vehicles will be considered only for half connections, there is no need to use the maximum number of parcels between hubs and line haul vehicles.

Since a service provider can rent vehicles daily, transportation costs can be calculated based on the capacity of dissimilar vehicles. To handle heterogeneous vehicles by extending the formulation of Čupić and Teodorović (2014), the number of needed vehicles of each kind will be considered by Relations (17) – (20). Relations (17) and (18) calculate the number of required vehicles in route k, and Relations (19) and (20) similarly calculate the number of required vehicles in line-hauls between hubs i and j. Consider Q_1, Q_2,…, Q_q as the capacity of q different vehicle sizes. When the volume of the collected parcels is less than Q_1, the operator uses the smallest kind of vehicle. If the volume of the parcels is greater than Q_{v-1} $(v < q)$ but less than Q_v, the operator uses the v-th kind of vehicle. However, if the volume of the collected parcels is greater than Q_q, the operator uses as many large vehicles as needed or as many large vehicles plus only one of the v-th $(v < q)$ kind of vehicles.

To better understand Relations (17) – (20), consider that there are three different vehicles with different costs and capacities used in the door-to-door services of Iran. Let Q_1, Q_2, and Q_3 be the capacity of the small, medium, and large vehicles. The related relations of the needed vehicles for the first route of this system can be computed by:

$$VR_{11} = \begin{cases} 1 \quad \text{if} \quad MMR_1 \le Q_1 \quad \text{or} \quad (MMR_1 > Q_3 \text{ and } MMR_1 - Q_3 \left\lfloor {}^{MMR_1}/_{Q_3} \right\rfloor \le Q_1) \\ \quad 0 \quad \text{otherwise} \end{cases} \tag{28}$$

$$VR_{21} = \begin{cases} 1 \quad \text{if} \quad Q_1 \le MMR_1 \le Q_2 \quad \text{or} \quad (MMR_1 > Q_3 \text{ and } Q_1 < MMR_1 - Q_3 \left\lfloor {}^{MMR_1}/_{Q_3} \right\rfloor \le Q_2) \\ \quad 0 \quad \text{otherwise} \end{cases} \tag{29}$$

$$VR_{31} = \begin{cases} 0 & \text{if} \quad MMR_1 \le Q_2 \\ \left\lfloor {}^{MMR_1}/_{Q_3} \right\rfloor & \text{if} \quad MMR_1 > Q_2 \text{ and } MMR_1 - Q_3 \left\lfloor {}^{MMR_1}/_{Q_3} \right\rfloor \le Q_2 \\ \left\lfloor {}^{MMR_1}/_{Q_3} \right\rfloor + 1 & \text{if} \quad D > Q_2 \text{ and } MMR_1 - Q_3 \left\lfloor {}^{MMR_1}/_{Q_3} \right\rfloor > Q_2 \end{cases} \tag{30}$$

When the collected parcel volume is less than Q_1, the operator uses a small vehicle. The operator uses a medium vehicle if the volume is greater than Q sub 1 but less than Q sub 2. Similarly, the operator uses a large vehicle if the needed capacity for parcels is more than Q sub 2 but less than Q sub 3. However, suppose the collected parcel volume is greater than Q sub 3. In that case, the operator uses as many large vehicles as needed or as many large vehicles plus one small or medium vehicle. The rational reason for using such a formulation for heterogeneous vehicles is that the rental cost of vehicles will not grow as fast as their capacities.

Finally, Constraints **Error! Reference source not found.** to **Error! Reference source not found.** determine the type of decision variables. Although Model I in this form has some non-linear combinations, with some changes in the following sub-section, the MINLP model can be changed to an MIP one.

Linearization

Relation (1) and Constraints (14), (16) contain two variables multiplied, at least one of which is binary. Consider binary variable S and positive variable T ($0 \leq \mathrm{T} \leq \bar{\mathrm{t}}$) and the product of these two variables as $\mathrm{U} = \mathrm{S}.\mathrm{T}$, we can easily linearize U as follows (Wolsey, 1998):

$$\mathrm{U} \leq \mathrm{S}\,.\bar{\mathrm{t}} \tag{31}$$

$$\mathrm{U} \leq \mathrm{T} \tag{32}$$

$$\mathrm{U} \geq \mathrm{T} - (1 - \mathrm{S})\bar{\mathrm{t}} \tag{33}$$

$$\mathrm{U} \geq 0 \tag{34}$$

The model minimizes the total cost of the system, and the amount of $\mathrm{VR}_{\mathrm{vk}}$ is totally dependent to $\mathrm{MMR}_{\mathrm{k}}$. Without loss of generality, Constraint (13) can easily be linearized by:

$$\mathrm{MMR}_{\mathrm{k}} \geq \ \mathrm{MRP}_{\mathrm{k}} \quad \forall\mathrm{k}, \tag{35}$$

$$\mathrm{MMR}_{\mathrm{k}} \geq \ \mathrm{MRD}_{\mathrm{k}} \quad \forall\mathrm{k}, \tag{36}$$

Relations (17) – (20) are also non-linear. We first provide linearization of (17) and (18) for the k-th route as follows, and then it can easily be generalized for Relations (19) and (20) of line haul vehicles.

Lemma. Constraints

$$\mathrm{VR}_{\mathrm{qk}} = \mathrm{VR}'_{\mathrm{qk}} + \mathrm{VR}''_{\mathrm{qk}} \quad \forall\mathrm{k} \tag{37}$$

$$\sum_{\mathrm{v}\neq\mathrm{q}} \mathrm{VR}_{\mathrm{vk}} + \mathrm{VR}''_{\mathrm{qk}} = 1 \quad \forall\mathrm{k} \tag{38}$$

$$\mathrm{MMR}_{\mathrm{k}} \leq \sum_{\mathrm{v}\neq\mathrm{q}} \mathrm{Q}_{\mathrm{v}}.\mathrm{VR}_{\mathrm{vk}} + \mathrm{Q}_{\mathrm{q}}.\mathrm{VR}'_{\mathrm{qk}} + \mathrm{Q}_{\mathrm{q}}.\mathrm{VR}''_{\mathrm{qk}} \quad \forall\mathrm{k} \tag{39}$$

$$\mathrm{MMR}_{\mathrm{k}} \geq \sum_{\mathrm{v}\neq\mathrm{q}-1,\mathrm{q}} \mathrm{Q}_{\mathrm{q}}.\mathrm{VR}_{(\mathrm{v}+1)\mathrm{k}} + \mathrm{Q}_{\mathrm{q}-1}.\mathrm{VR}''_{\mathrm{qk}} + \mathrm{Q}_{\mathrm{q}}.\mathrm{VR}'_{\mathrm{qk}} \quad \forall\mathrm{k} \tag{40}$$

where $\mathrm{VR}'_{\mathrm{qk}}$ and $\mathrm{VR}''_{\mathrm{qk}}$ are integer and binary variables, respectively. Correctly linearize Relations (17) and (18).

Proof. Depending on the values of $\mathrm{MMR}_{\mathrm{k}}$, q different cases (i.e., number of various kinds of vehicles) may occur as follows:

Case 1: $mQ_q < MMR_k \leq mQ_q + Q_1$ $(m = 0,1,2, \ldots)$, which means $MMR_k = mQ_q + a$ and $0 < a \leq Q_1$. Then, an operator needs one vehicle of the smallest kind and m vehicles of the largest one (q). Clearly, in this case $MMR_k \leq Q_1.VR_{1k} + Q_q VR'_{qk}$ and $MMR_k \geq Q_q.VR'_{qk}$. Hence, $VR_{1k} = 1$, and $VR'_{qk} = \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor$ and other variables are equal to 0.

Case v $(v < q)$: $mQ_q + Q_{v-1} < MMR_k \leq mQ_q + Q_v$ $(m = 0,1,2, \ldots \text{and } Q_0 = 0)$, which means $MMR_k = mQ_q + a$ and $Q_{v-1} < a \leq Q_v$. Then, an operator needs one vehicle of the v-th kind and m vehicles of the largest kind. Obviously, in this case, $MMR_k \leq Q_v.VR_{vk} + Q_q VR'_{qk}$ and $MMR_k \geq Q_q.VR'_{qk}$. Hence, $VR_{vk} = 1$ and $VR'_{qk} = \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor$ and other variables are equal 0.

Case q: $mQ_q + Q_{q-1} < MMR_k \leq (m+1)Q_q$ $(m = 0,1,2, \ldots)$, which means $MMR_k = mQ_q + a$ and $Q_{q-1} < a \leq Q_q$. Then, an operator needs (m+1) vehicles of the largest one. In this case, $MMR_k \leq Q_q.VR'_{qk} + Q_q.VR''_{qk}$ and $MMR_k \geq Q_{q-1}.VR''_{qk} + Q_q.VR'_{qk}$. Hence, $VR''_{qk} = 1$ and $VR'_{qk} = \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor$ and other variables are equal to 0.

Therefore, by using binary variables of VR_{vk} $(v = 1,2, \ldots, q-1)$ and VR''_{qk} and an integer variable of VR'_{qk}, all q possible cases are guaranteed by Constraints (38) to (40). To better understand the proposed cases and results in each case, please see Figure 2 .

Case 1 $\left(mQ_q < MMR_k \leq mQ_q + Q_1\right)$, then we have:

$$VR_{1k} = 1, \ldots, VR_{vk} = 0, \ldots, VR_{qk} = \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor \cong (VR'_{qk} = \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor \text{and } VR''_{qk} = 0)\,;$$

⋮

Case $v(v < q)$ $(mQ_q + Q_{v-1} < MMR_k \leq mQ_q + Q_v)$ Then

$$VR_{1k} = 0, \ldots, VR_{vk} = 1, \ldots, VR_{qk} = \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor \cong (VR'_{qk} = \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor \text{and } VR''_{qk} = 0)\,;$$

⋮

Case q $(mQ_q + Q_{q-1} < MMR_k \leq (m+1)Q_q)$, then we have:

$$VR_{1k} = 0, \ldots, VR_{(q-1)k} = 0, VR_{qk} = \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor + 1 \cong (VR'_{qk} = \left\lfloor {}^{MMR_k}/_{Q_q} \right\rfloor \text{and } VK''_{qk} = 1)\,;$$

Figure 2. Number of heterogeneous vehicles on the k-th route

Finally, by linearizing the related constraints and objective function, we modify Model I to the MIP form of Model II.

Model II:

Objective function (1)

s.t.

(2) – (14), (16), (21) – (23), (26), (27), (35), (36), and (37) – (40).

$$VH_{qij} = VH'_{qij} + VH''_{qij} \quad \forall i,j \tag{41}$$

$$\sum_{v \neq q} VH_{vij} + VH''_{qij} = 1 \quad \forall i,j \tag{42}$$

$$MMH_{ij} \leq \sum_{v \neq q} Q_v . VH_{vij} + Q_q . VH'_{qij} + Q_q . VH''_{qij} \quad \forall i,j \tag{43}$$

$$MMH_{ij} \geq \sum_{v \neq q-1,q} Q_q . VH_{(v+1)ij} + Q_{q-1} . VH''_{qij} + Q_q . VH'_{qij} \quad \forall i,j \tag{44}$$

$$VR_{vk} = \{0,1\} \quad \forall v \neq q\,, k, \tag{45}$$

$$VR'_{qk} = \{0,1\} \quad \forall k, \tag{46}$$

$$VR''_{qk} \text{ is integer} \quad \forall k, \tag{47}$$

$$VH_{vij} = \{0,1\} \quad \forall v \neq q\,, i, j, \tag{48}$$

$$VH'_{qij} = \{0,1\} \quad \forall i,j, \tag{49}$$

$$VH''_{qij} \text{ is integer} \quad \forall i,j. \tag{50}$$

Although Model II is transformed to the MIP formulation, such integrated location and routing problems are difficult to solve, especially for large practical problems (Čupić & Teodorović, 2014). Parallel, heuristic, and meta-heuristic approaches promise the highest quality of solutions in reasonable time periods.

Solution method

A new two-stage hybrid method is proposed to set hub locations, node allocations, and multi-path routings in the first stage, and improve allocation and routings in the second stage.

Solution representation

Four different strings are designed to convert the solution network into comprehensible code, as shown in Figure 3. If the elements of the strings represent the number of nodes, the first string determines the selected hub. In the first string, at first, all elements are set to zero. However, if some elements take the value of one, it means the related nodes are chosen as the hubs, and the others are non-hub nodes. The second string represents the allocation of nodes to the hubs, and its elements can only take the number of nodes that are selected as hubs. In this string, a hub node should take its own number, to show it (as a node) is allocated to itself (as a hub). The third and fourth strings belong to the path routing, in which the third string determines the allocated route of each node, while the fourth one shows the sequence of each node in its related route. As hub nodes always take the start positions and may start more than one route, they take zero in the third and fourth string. In Figure 3, an example with 10 nodes and 2 hubs is used to show the solution representation method. In this example, nodes 3 and 7 are chosen as the hubs, so their related elements in the first string are equal to 1. Nodes 1, 2, 4, and 9 are allocated to the hub node 3, while nodes 5, 6, 8, and 10 are allocated to the hub node 7. Five different routes are formed to link non-hub nodes to the hubs; for example, nodes 1 and 2 are linked to the hub node 3 by route 1, in which nodes 2 and 1 are in the first and second positions, respectively.

	1	2	3	4	5	6	7	8	9	10
Hub	0	0	1	0	0	0	1	0	0	0
Allocation	3	3	3	3	7	7	7	7	3	7
Route	1	1	0	2	3	3	0	4	5	4
Sequence	2	1	0	1	1	2	0	2	1	1

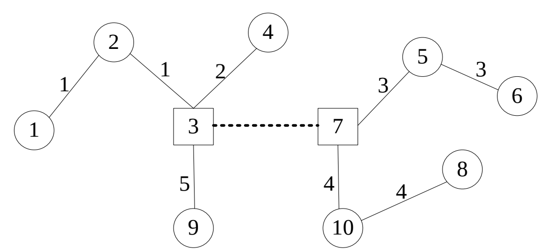

Figure 3. Example with 10 nodes and 2 hubs

Generating initial feasible solutions

Different items should be determined to provide a feasible solution (i.e., hub combination, allocation of nodes to hubs, and multi-path routing) to fill the four mentioned strings.

a) Generating hub combination

Decreasing the complexity of the proposed method, the number of hubs (Nhub) in each run is fixed at about the following two conditions:

- A feasible hub combination consists of the minimum number of hubs, so each node is covered by at least one hub in less than TR2.

- Connecting all hubs, the time periods/distances between selected hubs should be less than TH.

b) Allocating nodes to hubs

Offering an initial solution, each node is allocated to its nearest hub. In this way, the method can provide a rational and feasible solution. Preserving the final solution's feasibility is unnecessary, as it is checked when choosing the hub combination.

c) Generating initial multi-path routing

More accurate calculation of the parcel delivery system expenses necessitates using routing methods within the location part. Since numerous distinct path-routes exist for each hub, two different methods have been designed based on forward and backward scheduling. In the forward routing, nodes are allocated to the routes based on their closeness to the hubs or the latest nodes of the current route, a greedy approach. In the backward routing, the last node of the route is first allocated based on the farthest distance to the hub, and other nodes can add to the current route only when they do not violate the route's feasibility. For each solution, one routing method can be chosen randomly based on the probability of using forward against backward (PrForward).

Generating neighborhood solutions

Since the presented method needs to improve the current solution in different facets of hub location, node allocation, and multi-path routing, three different neighbor procedures are introduced to cover these aspects.

a) Hub location neighborhood

If N shows the total number of nodes and the number of hubs in each step is fixed, a feasible hub location solution consists of Nhub hubs and N-Nhub non-hub nodes. The NH hub is randomly removed from the solution to generate a new solution, and the NH non-hub node will be selected as a hub to enter the new solution. However, the procedure preserves the feasibility of the new solution by checking the time period constraints of the maximum allowable time/distance from hubs to hubs and nodes to hubs. The value of variable NH can be a fixed number or changed during the algorithm's process.

To generate a hub location neighborhood of Figure 3, the hub node 3 is randomly removed from the hub list, and node 4 is randomly selected as a new hub. The effect of this change on the solution and network is illustrated in Figure 4.

	1	2	3	4	5	6	7	8	9	10
Hub	0	0	0	1	0	0	1	0	0	0
Allocation	4	4	4	4	7	7	7	7	4	7
Route	1	1	2	0	3	3	0	4	2	4
Sequence	2	1	1	0	1	2	0	2	2	1

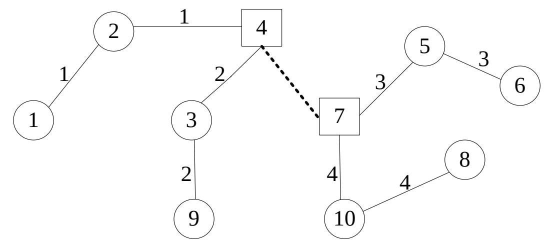

Figure 4. Hub location neighborhood of Figure 2

b) Node allocation neighborhood

Nodes are allocated to their nearest hubs in initial solutions; however, this approach cannot guarantee finding the best solutions. In the proposed neighborhood procedure, NN nodes are randomly selected and allocated to other hubs. However, the algorithm preserves the feasibility of the solution by checking the time windows constraints between the newly allocated nodes and their new hubs. The value of variable NN can be a fixed number or changed during the algorithm's process. By randomly allocating node 4 to hub node 7, a node allocation neighborhood of Figure 3 is represented in Figure 5.

	1	2	3	4	5	6	7	8	9	10
Hub	0	0	1	0	0	0	1	0	0	0
Allocation	3	3	3	7	7	7	7	7	3	7
Route	1	1	0	2	3	3	0	4	5	4
Sequence	2	1	0	1	1	2	0	2	1	1

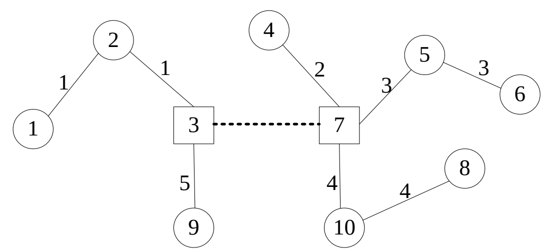

Figure 5. Node allocation neighborhood of Figure 2

c) Multi-path routing neighborhood

Initial multi-path routing is composed of two procedures, forward and backward routing; however, the routing part may need to be improved. In this neighborhood procedure, NR nodes are selected and randomly inserted in another route of their or other hubs, if possible. The feasibility of the final solution should be preserved, as routes cannot violate extended time. The value of variable NR can be a fixed number or can be changed based on the solution space. In Figure 6, a multi-path routing neighborhood of Figure 3 is generated by randomly changing the route of node 8 from 3 to 4.

	1	2	3	4	5	6	7	8	9	10
Hub	0	0	1	0	0	0	1	0	0	0
Allocation	3	3	3	3	7	7	7	7	3	7
Route	1	1	0	2	3	3	0	3	5	4
Sequence	2	1	0	1	1	2	0	3	1	1

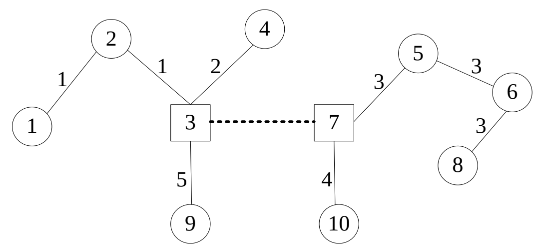

Figure 6. Multi-path routing neighborhood of Figure 2

Two-stage method

In the first stage, the method finds the best combination of hubs in the system and allocates all nodes to the selected hubs, while multi-path routing is determined based on the forward or backward method. The important task of simultaneously determining hub location and node allocation needs an approach (e.g., artificial bee colony (ABC)), which can explore and exploit the solution space at the same time. An ABC algorithm that simulates the smart foraging behavior of a honeybee swarm has shown better performance in comparison with other population-based algorithms, with the advantages of using fewer control parameters (Karaboga & Akay, 2009). In the proposed ABC algorithm, three types of bees, namely employed bees, onlooker bees, and scout bees, are applied. The employed bees search for initial solutions based on the number of fixed hubs. By calculating the objective function value (OFV), the employed bees share their information with the onlooker bees. As the onlooker bees tend to select the solution with fewer OFVs, they calculate the fitness function as (51) for all solutions and choose the one with higher F_i.

$$F_i = \frac{1/OFV_i}{(\sum_i 1/OFV_i)/NE_i} \tag{51}$$

where NE is the number of employed bees.

This probabilistic selection can be performed based on the roulette wheel selection mechanism in a way that the greater the value of F_i, the greater the probability of searching (PSearch), the current solution space by onlooker bees. When a solution is not selected for searching by onlooker bees, it should be generated by a scout bee. In fact, a new solution can be produced by the hub neighborhood with slight changes to hub nodes or replaced by a randomly generated solution.

In the second stage, the method amends the routing and node allocation (if possible) by a Simulated Annealing (SA) method and two local search (LS) methods, as the method should exploit the current solutions and improve the final output.

SA is a probabilistic technique that imitates the condition of a hot solid, which is gradually cooled down to become a frozen one. The algorithm was independently proposed by Kirkpatrick et al. (1983) and Černý (1985) to discover or approximate the global optimum of an objective function, especially in a large search space. This algorithm starts from the output of the previous stage. If the new solution is better than the best solution found so far, the new solution substitutes the best solution. Afterwards, as mentioned before, two simple LS methods are considered. In the first LS, the method deliberates all routes and allocates all or part of them to the second nearest hub, if it is possible. In the second LS, the method looks at the extended routes to probably change allocated nodes of all or part of the hubs to decrease route time/distance. In both LS methods, a new solution will be accepted only when the changes can decrease the total expenses of the parcel delivery system.

In the proposed method, the ABC and SA algorithms stop when the maximum iterations conditions arrive or the algorithm cannot improve the best solution in pre-determined loops. The descriptive flow chart of the procedure is shown in Figure 7, where N-best is the best solution identified so far.

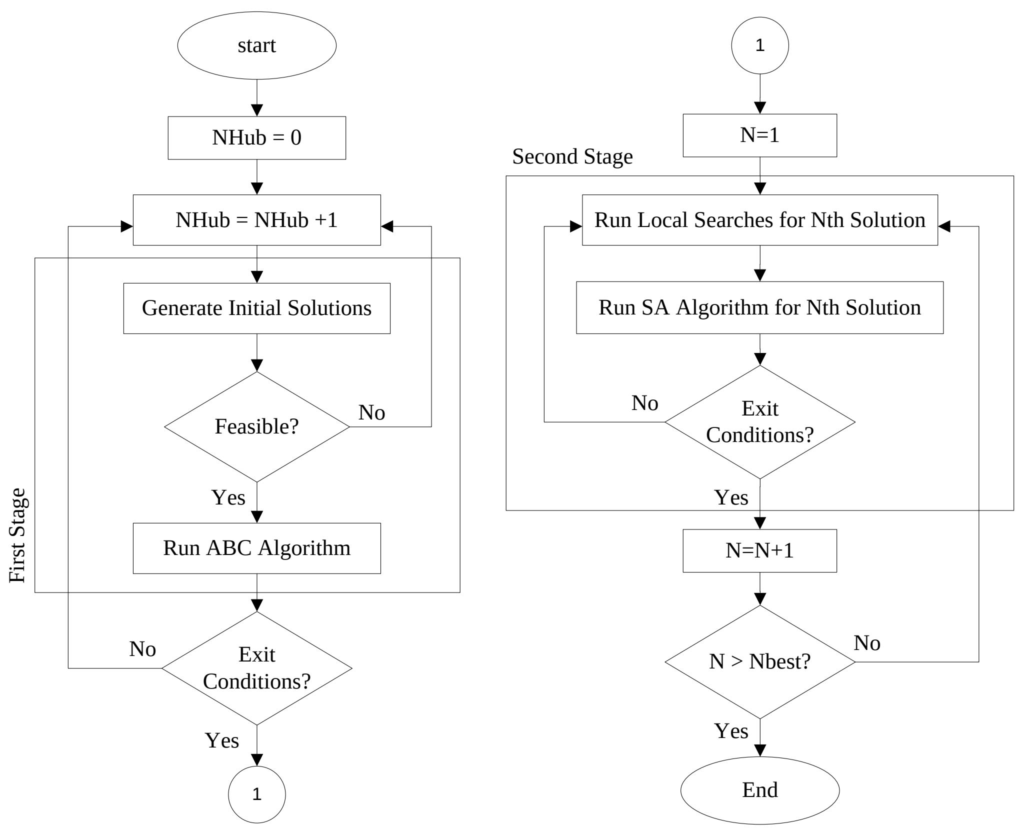

Figure 7. Flow chart of the proposed method

Results

Computational results

As emphasized earlier, the considered parcel delivery network is supposed to service a broad and sparse country of Iran. Contrary to the common practice of such logistic systems, incomplete roads and distant cities prevent forming tours between hubs and nodes (cities) and force the system provider to offer three different delivery time windows. To evaluate the proposed model and method, 75 test instances with five sizes of 10, 15, 20, 30, and 50 nodes in three area spaces of 750×750, 1000×1000, and 1250×1250 km^2 are created based on the real conditions. In fact, for each node number, 15 test instances in three area spaces are shaped (five instances in each area space).

To assess the performance of the proposed solution method, small-sized instances are solved optimally, and some instances in all sizes are solved by a SA-based method (Bahrami et al., 2016). The results of test instances and two practical cases of Iran, with 31 and 59 cities, are used for further examinations and sensitive analysis. It should be mentioned that the model is coded in GAMS 24.1 to solve with the CPLEX solver, and the proposed two-stage hybrid method is coded in MATLAB R2014a. The model and method run by Intel Core i5, 3.1 GHz compiler with 8 GB of RAM, in such a way that the CPLEX operates in the parallel processing mode; however, the MATLAB program is run in a single process.

All hubs should be connected to each other via direct links, and they transfer the parcels on a night shift to make use of low-traffic roads. While vehicles that travel in the daylight shift in routes with stopovers may encounter some circumstances, such as mechanical breakdown, traffic before and after cities, damaged roads, or prolonged loading and unloading time in each city. Therefore, the investor is convinced that such unpredictable conditions prevent route vehicles from reaching their maximum speed. So, the average speed of vehicles in the path routing part is considered 75 km/h; however, the average speed of line haul vehicles can exceed 85 km/h.

Consulting the company managers, two sets of the maximum distances of hubs are taken into account based on the distances that vehicles can travel on a day: TH=680 and TK1=600, or TH=510 and TK1=675. For both cases, TK2=2×TK1+150 is considered to support far and distant cities. It should be mentioned that the cost of vehicles will be similar in both cases, as they should be recruited for all day.

To set distances between nodes in the designed test instances, the p-norm distance of Minkowski is used as a Relation (52). When p=1, it calculates the city block distance, and with p=2, it computes the direct distance between two nodes. However, in the proposed problem, something between p=1 and p=2 can illustrate the real distance between two cities. As shown in Figure 8, p=1.3 can approximately simulate the practical and intended distances.

$$d = \left(\sum_{i=1}^{n} |a_i - b_i|^p\right)^{1/p} \tag{52}$$

where ai and bi are two n-dimensional points in the Cartesian coordinates.

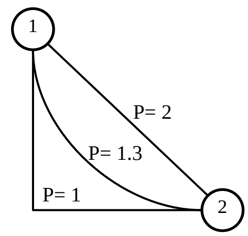

Figure 8. Three different p-norm distances of Minkowski

Three different vehicle types are considered with capacities of 2000, 3000, and 4000 kg, with the costs of 25, 35, and 40 monetary units for collecting and distributing in the routes, respectively, while they will be half price for line haul transportation ($\alpha = 0.5$). As the logistic company is new in the market and can only invest in some big cities, the number of bundled parcels between two cities is created randomly in range of [0, 2000] kg with some consideration; the volume of demand between about 50% of the cities is considered less than 100 kg, and for only 15% of cities the transferred demand volume is considered in range of [300, 2000] kg. This assumption tries to simulate the LTL strategy and to make use of all possible combinations of vehicle types. Furthermore, the fixed cost of establishing a hub is randomly generated in a range of [250, 300]. To prevent opening false routes with one city, the cost of establishing a route is considered 25 monetary units. Finally, delighting the customers, the company should be able to improve its 24-hour delivery services. To encourage a normal time window and provide additional finance for vehicles of extended time routes, a penalty cost of 0.2 is considered for the parcels collected or delivered on the extended routes.

Parameter settings

The effectiveness of the meta-heuristics depends entirely on the selected values of their parameters. When the number of parameters rises, testing all possible cases is not economical. Controlling the time and cost of tuning the factors of the solution methods, several experimental design techniques are suggested to lessen the number of experiments (Bement, 1989). In this paper, the Taguchi method, which has been efficiently applied to optimization problems (Vahdani et al., 2012), is used for the parameter setting of two proposed meta-heuristics.

Comparisons

Two different comparisons are taken into account to evaluate the proposed solution methods; the first one is the comparison of the results of the proposed scheme with those of the optimal techniques, and the second one is the assessment of the method outcomes in contrast with those of the SA-based method presented by Bahrami et al. (2016).

Since the proposed model solved by CPLEX is NP-hard, only small and medium test instances with 10, 15, and 20 nodes are chosen for the first comparison, and the CPU time limits of 3600, 7200, and 10800 seconds are considered, respectively. Except for some 15 and 20-node test instances in the area space of 750×750 km2, other test instances with 15 and 20 nodes are not able to reach a reasonable result; thus, only 10-node test instances are reported in Table 3. It should be mentioned that the objective function value (OFV) is in monetary units and the CPU time is in seconds. Besides, each instance is illustrated by three numbers (e.g., 10-2-5), which show the node number, area space (1 for 750×750 km2, 2 for 1000×1000 km2, and 3 for 1250×1250 km2), and number, respectively.

Solving 30 instances (i.e., 15 test instances in two different THs), the CPLEX proves the optimality of 18 cases and comes across the best possible output of 11 cases, but does not report any solution for one instance. The solution method finds equivalent results for 25 cases; however, in four cases, the outcomes are not as good as those obtained by the CPLEX. The underlined digits in Table 3**Error! Reference source not found.** show the instances in which the method cannot optimally determine their solutions. The results indicate less than 1% difference between the solution method and CPLEX output, while for all instances, the proposed method finds the result with less CPU time and, on average, 2600% faster than that of CPLEX.

Table 3. Comparison of the solution method with CPLEX in 10-node test instances

Area	TH	Instances	GAMS		The Method		Difference	
			OFV	Time	OFV	Time	%OFV	%Time
750×750	680	10-1-1	584.0	2966.8	584.0	98.24	0.000	29.20
		10-1-2	790.0	3600	790.0	97.30	0.000	36.00
		10-1-3	533.0	267.9	533.0	95.44	0.000	1.81
		10-1-4	551.0	3600	551.0	114.03	0.000	30.57
		10-1-5	669.0	3348.5	669.0	109.85	0.000	29.48
	510	10-1-1	559.0	2315.7	559.0	79.07	0.000	28.29
		10-1-2	529.0	1705.7	529.0	95.84	0.000	16.80
		10-1-3	508.0	723.3	508.0	49.29	0.000	13.67
		10-1-4	517.0	502.3	517.0	68.20	0.000	6.37
		10-1-5	618.0	1541.6	618.0	82.21	0.000	17.75
1000×1000	680	10-2-1	1010.0	1412.6	1010.0	101.12	0.000	12.97
		10-2-2	854.0	3600	854.0	92.13	0.000	38.07
		10-2-3	1148.0	1181.5	1148.0	42.37	0.000	26.88
		10-2-4	1222.0	3600	1222.0	86.50	0.000	40.62
		10-2-5	1555.0	3600	1555.0	87.77	0.000	40.02
	510	10-2-1	1792.4	281.8	<u>2036.4</u>	80.36	0.136	2.51
		10-2-2	597.0	644.7	597.0	67.15	0.000	8.60

		10-2-3	548.0	177.0	548.0	61.06	0.000	1.90
		10-2-4	944.0	371.6	944.0	60.57	0.000	5.14
		10-2-5	891.0	1487.0	891.0	60.28	0.000	23.67
1250×1250	680	10-3-1	2347.4	3600	2347.4	86.87	0.000	40.44
		10-3-2	1923.8	3600	1923.8	103.54	0.000	33.77
		10-3-3	869.0	1607.5	869.0	86.08	0.000	17.68
		10-3-4	1241.0	1395.6	1241.0	86.58	0.000	15.12
		10-3-5	2351.8	643.3	2358.8	99.67	0.003	5.45
	510	10-3-1	2832.6	3600	3085.1	60.27	0.089	58.73
		10-3-2	1796.4	3600	1838.4	60.41	0.023	58.59
		10-3-3	861.0	3600	861.0	60.99	0.000	58.03
		10-3-4	***	1029.6	1653.2	59.97	***	***
		10-3-5	2600.5	3600	2600.5	60.90	0.000	58.11

Although the method cannot determine the optimal solution of four instances in one run, it finds the optimal results of all 10-node test instances with some changes. One or two changes, including a few modifications in some parameters of NN, NH, and NR, among others, and especially increasing N-best for two instances of 10-2-1 (TH=510) and 10-3-1 (TH=510), suggest better searching of the solution space.

Comparisons of the proposed method in all node sizes are done with the SA-based method (Bahrami et al., 2016). Table 4 illustrates the results of solving the first examples of all nodes in all area spaces. As shown in this table, there are no significant differences between the results of the two methods in solving small-sized test instances. The technique delivers much better outputs by increasing the number of nodes and expanding the area spaces, except in four cases. In 50-node test instances, the proposed method provides better results by an average of 4%. Generally, the solution method can form its results by an average of 11% faster than the SA-based method outcomes.

Table 4. Comparison of the solution method with the SA-based method in all sizes

Area	TH	Instances	SA based		Proposed Method		Difference	
			OFV	Time	OFV	Time	%OFV	%Time
750×750	510	10-1-1	559.0	81.9	559.0	79.1	0.000	0.035
		15-1-1	812.0	252.4	812.0	223.3	0.000	0.115
		20-1-1	1067.0	430.9	1103.0	355.8	-0.034	0.174
		30-1-1	2630.0	903.7	2610.0	776.5	0.008	0.141
		50-1-1	5380.0	1392.6	5243.0	1027.0	0.025	0.263
	680	10-1-1	584.0	68.0	584.0	98.2	0.000	-0.445
		15-1-1	829.0	208.3	829.0	215.5	0.000	-0.034
		20-1-1	1138.0	401.6	1140.0	365.5	-0.002	0.090
		30-1-1	2870.0	693.3	2785.0	558.6	0.030	0.194
		50-1-1	5612.0	1229.9	5566.0	1186.0	0.008	0.036
1000×1000	510	10-2-1	1760.0	81.7	1792.4	80.4	-0.018	0.016
		15-2-1	1782.0	245.5	1780.0	205.6	0.001	0.162

		20-2-1	1659.0	389.2	1659.0	397.6	0.000	-0.022
		30-2-1	2893.0	583.9	2613.5	565.0	0.097	0.032
		50-2-1	6950.5	1363.5	6560.5	1231.0	0.056	0.097
	680	10-2-1	1010.0	93.7	1010.0	101.1	0.000	-0.079
		15-2-1	1206.0	164.9	1206.0	214.8	0.000	-0.303
		20-2-1	1735.0	502.0	1730.5	430.1	0.003	0.143
		30-2-1	2744.5	624.5	2744.5	509.1	0.000	0.185
		50-2-1	7350.0	1295.7	7014.0	1066.0	0.046	0.177
1250×1250	510	10-3-1	3002.3	73.8	2832.6	60.3	0.057	0.184
		15-3-1	2447.2	231.9	2447.2	207.7	0.000	0.104
		20-3-1	6327.8	533.6	5988.5	406.4	0.054	0.238
		30-3-1	6234.0	778.0	5743.1	591.0	0.079	0.240
		50-3-1	24560.1	1429.3	23204.0	1214.0	0.055	0.151
	680	10-3-1	2352.0	81.7	2347.4	86.9	0.002	-0.064
		15-3-1	4498.8	245.5	4518.3	232.5	-0.004	0.053
		20-3-1	6435.6	389.2	6283.0	422.5	0.024	-0.086
		30-3-1	6349.1	583.9	5823.6	566.4	0.083	0.030
		50-3-1	15890.0	1363.5	15593.0	1743.0	0.019	-0.278

Sensitive analysis

Analyzing the effectiveness of the proposed method, all 75 test instances in two THs are solved three times. The average of the results is summarized in Table 5. In this table, S1 and S2 show the best OFV of the first and second stage, respectively, and Diff means the difference between the OFV of the best solution in the first stage minus the OFV of the second stage. Since N-best solutions of the first stage enter the second stage, the OFV of the best solution in the first stage is not necessarily the best OFV of S1. Nhub and Nroute mean the number of hubs and routes, respectively. Ncon means the number of nodes' relationships that belong to each delivery time window of 24, 48, and 72 hours. Finally, V-Route and V-Line Haul illustrate the number of vehicles in three capacities: L for large vehicles, M for medium vehicles, and S for small cars in routes and between hubs, respectively. In Table 5, instances are shown with two digits, the first digit illustrates the node number, and the second one shows the area spaces.

Table 5. Result of the solution method solving all test instances

Instances	TH	OFV			Nhub	Nroute	Ncon			V-Rout			V-Line Haul			Time	
		S1	S2	Diff			24	48	72	L	M	S	L	M	S	S1	S2
10-1	510	564.0	546.2	26.20	1.0	3.6	45.0	0.0	0.0	3.4	1.6	0.4	0.0	0.0	0.0	38.1	36.8
10-1	680	637.8	625.4	16.60	1.2	4.0	39.0	4.8	1.2	3.2	1.0	2.0	0.0	0.4	0.0	64.4	38.6
10-2	510	1006.9	954.5	52.40	1.6	5.2	43.2	1.8	0.0	2.0	0.8	3.0	0.5	0.5	0.0	38.0	27.9
10-2	680	1321.5	1157.8	163.68	2.0	5.2	35.4	8.4	1.2	2.0	1.0	2.8	0.8	0.4	1.6	49.8	32.1
10-3	510	2073.2	1957.1	67.38	1.8	6.4	31.4	12.8	0.8	1.8	1.8	3.0	1.0	0.6	0.2	35.9	24.6

10-3	680	1748.0	1748.0	0.00	2.2	5.2	34.6	10.0	0.4	1.4	1.6	2.8	1.0	1.0	1.2	57.1	35.4
15-1	510	871.2	837.6	50.00	1.0	4.8	105.0	0.0	0.0	10.0	0.6	1.4	0.0	0.0	0.0	84.0	144.5
15-1	680	912.6	864.0	61.80	1.0	5.6	105.0	0.0	0.0	9.6	0.8	2.0	0.0	0.0	0.0	93.4	143.8
15-2	510	1179.2	1153.8	31.00	1.6	6.8	102.2	2.8	0.0	6.2	1.8	3.0	2.4	0.0	0.8	74.9	146.8
15-2	680	1088.8	1072.6	16.40	1.6	6.0	64.2	11.6	29.2	6.0	2.0	3.0	2.6	0.4	0.2	88.6	151.0
15-3	510	3436.9	3400.0	160.22	1.8	7.4	80.2	21.6	3.2	7.6	1.6	3.6	2.8	0.8	0.8	68.1	155.3
15-3	680	3783.6	3779.6	4.00	2.8	7.4	76.8	25.6	2.6	6.6	1.8	3.6	3.6	0.6	4.2	72.4	158.0
20-1	510	1237.4	1182.4	85.00	1.0	6.2	190.0	0.0	0.0	15.6	2.0	2.4	0.0	0.0	0.0	174.7	210.0
20-1	680	1322.5	1275.9	50.60	1.2	7.0	190.0	0.0	0.0	15.8	2.4	1.4	2.0	0.0	0.0	196.3	205.9
20-2	510	1595.0	1569.5	46.30	1.8	7.8	190.0	0.0	0.0	16.6	1.0	3.8	5.0	0.2	0.8	194.4	190.7
20-2	680	1728.4	1684.7	43.70	2.0	8.4	190.0	0.0	0.0	16.0	2.2	3.4	6.6	0.6	0.6	213.6	195.3
20-3	510	5703.3	5612.6	156.76	2.2	9.4	143.2	44.2	2.6	14.2	2.6	3.4	8.2	1.2	0.8	182.2	203.3
20-3	680	4459.5	4428.0	51.80	3.0	9.6	154.4	33.2	2.4	13.0	3.0	3.2	9.4	2.2	1.8	167.4	212.1
30-1	510	2350.5	2273.8	106.00	1.2	7.8	435.0	0.0	0.0	40.4	2.0	2.0	1.4	0.2	0.2	291.6	324.1
30-1	680	2489.3	2435.8	64.60	1.4	9.2	435.0	0.0	0.0	39.8	2.0	3.4	4.2	0.6	0.2	236.0	302.0
30-2	510	2955.0	2882.8	77.40	2.2	9.0	435.0	0.0	0.0	39.4	2.6	3.6	14.2	0.8	1.2	264.5	290.7
30-2	680	3067.7	2977.0	97.50	2.2	10.0	435.0	0.0	0.0	38.8	3.0	3.8	18.2	0.6	0.2	284.9	282.9
30-3	510	9484.1	9408.1	90.00	2.8	12.6	362.8	69.0	3.2	36.2	4.0	5.4	23.0	1.2	2.2	388.5	265.8
30-3	680	8868.4	8838.0	45.70	3.4	13.4	368.0	64.4	2.6	36.0	2.8	7.2	26.2	2.0	2.2	390.9	265.6
50-1	510	5419.6	5295.0	150.80	1.0	9.4	1225.0	0.0	0.0	116.6	2.6	1.8	0.0	0.0	0.0	572.7	514.5
50-1	680	5595.8	5463.0	158.00	1.4	11.0	1225.0	0.0	0.0	115.8	1.6	3.2	2.0	0.2	0.6	550.2	530.6
50-2	510	6607.1	6523.4	154.20	2.0	15.2	1225.0	0.0	0.0	116.4	4.6	4.2	32.2	0.4	1.0	675.3	477.4
50-2	680	7196.9	7069.8	140.80	2.4	17.2	1225.0	0.0	0.0	114.6	3.8	6.0	56.0	0.2	1.8	673.6	504.5
50-3	510	18318.8	18162.0	210.20	3.4	17.6	1100.2	122.2	2.6	117.2	4.4	6.2	71.2	1.4	4.2	1281.0	502.4
50-3	680	16065.7	15930.7	171.80	4.0	18.4	1128.2	95.6	1.2	112.4	6.8	7.0	85.6	3.0	7.2	934.0	492.3

It can be seen from Table 5 that the second stage substantially impacts the results, especially on medium and large test instances with bigger area spaces. On average, the OFV is improved in the second stage by 1.6%. This cost saving can be critical since almost all expenses of the hub establishment, node connections, parcel delivery, and even route establishment and hiring of the vehicles are inevitable. In Figure 9-a, the mean percent of improvement (Diff) in the second stage is illustrated for all nodes.

Comparing the results of the first and second stages reveals more information about the method's behavior. On average, Nhub decreased by 1% while Nroute was reduced by 11% (in 45% of instances); even in one case, this reduction was 23%. In Figure 11-b, the mean value of Nroute in the first and second stages is illustrated for different nodes.

Analyzing time windows of the results reveals that the second stage intensifies the number of 24-hour transportations by reducing the number of 48 and 72-hour transports. In Figure 11-c, the mean value of changes in 24-hour transportation in the second stage is illustrated for all nodes. On average, 24-hour transportation increased by 4%.

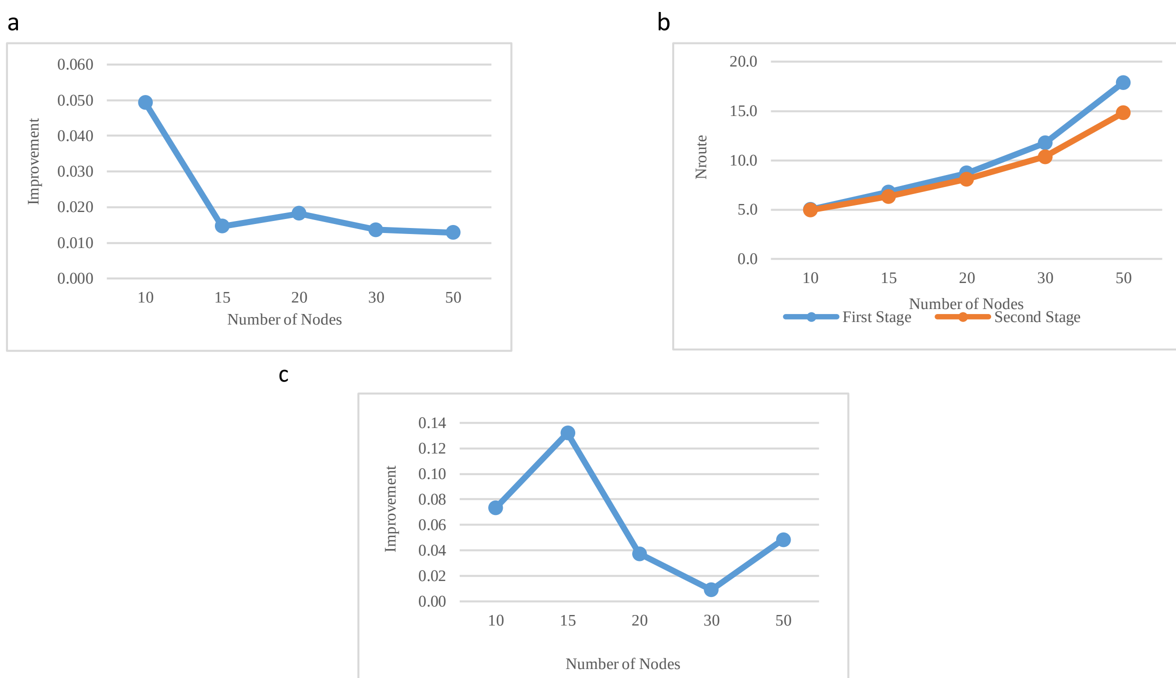

Figure 9: a) Mean percent of improvement in the second stage. b) Mean value of Nroute in the first and second stages. c) Changes in 24-hour transportation in the second stage

The effect of the second stage can be seen even in the combination of the vehicle fleet, as the number of large vehicles on the routes and line hauls decreased by 1% and 2%, respectively. Explicit trends in small and medium vehicle sizes cannot be seen by running the second stage.

It is worth mentioning that preserving the N-best solution of the first stage is a crucial factor, and in 42% of instances, the best final output is not yielded from its best solution of the first stage.

To consider the influence of the vast and dispersed country on the results, test instances are created in three areas of 750×750, 1000×1000, and 1250×1250 km². Additionally, two cases are considered for maximum distances of 510 and 680 km line hauls. In Figure 10, the concurrent influence of these two factors can be realized on the means of OFVs for different node numbers.

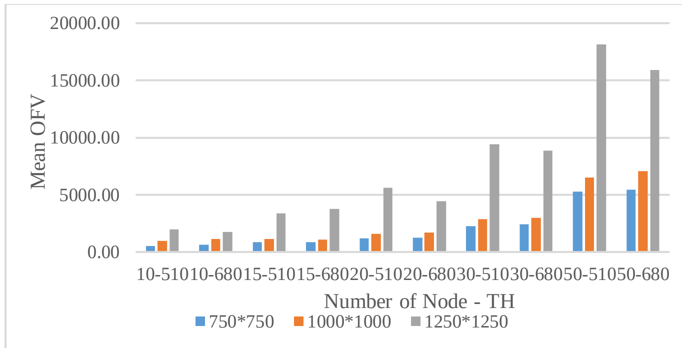

Figure 10. Impact of the node number and TH on the mean of OFVs

As you can see in Figure 10, by growing the area space and node numbers, the OFV enormously increases, especially in 1250×1250 km2 area space instances. However, the differences between the OFV of 750×750 and 1000×1000 km² are not substantial. Besides, in 750×750 and 1000×1000 km2 cases, TH=680 cases increase the OFV, but in 1250×1250 km2 area space, TH=510 has a bigger OFV than the second case. By expanding the area spaces, the OFVs of the instances increase. The reason is that the distances between the cities extend, and we need to increase the number of hubs, routes, and extended routes, which incur penalty costs. To better understand the reason for such behavior, it is necessary to consider the number of mean hubs and routes in Figure 11.

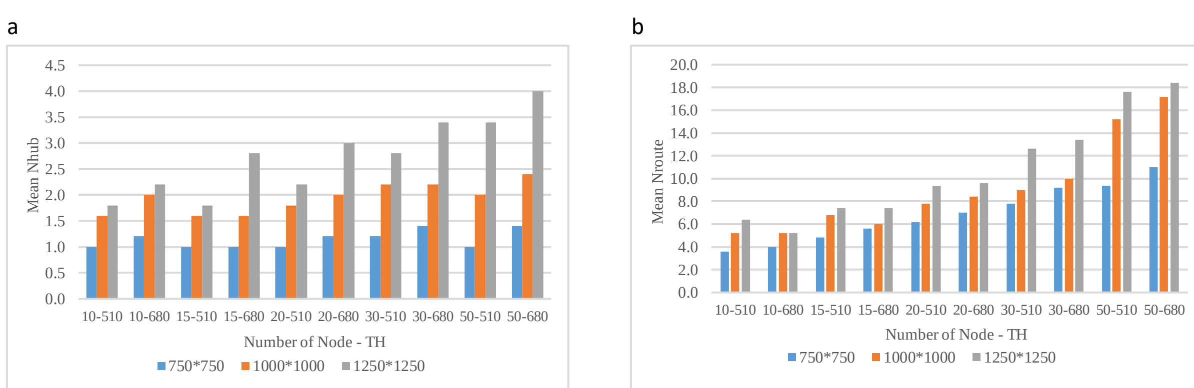

Figure 11 a) Impact of the node number and TH on mean Nhub. b) Impact of the node number and TH on mean Nroute.

TH=510 needs fewer hubs and routes in the same number of nodes (except 10 and 15 nodes in a 1000×1000 km² area space) than TH=680. It makes the model in 750×750 and 1000×1000 km² to reduce the system's expenses for test instances of TH=510. However, in the 1250×1250 km2 area, the closeness of the hubs in TH=510 makes the routes longer. So, test instances with TH=680 have less extended routes than instances with TH=510, which can reduce the OFV of TH=680 in all cases. In Figure 12, the mean ratio of the extended routes to regular routes in the 1250×1250 km2 area is illustrated as a decreasing trend. Except for 15 node cases, all cases of TH=680 have less extended routes. It is evident that for a bigger area space of 1250×1250 km2, longer distances between hubs provide better circumstances to cut expenses and cover the whole area and all nodes.

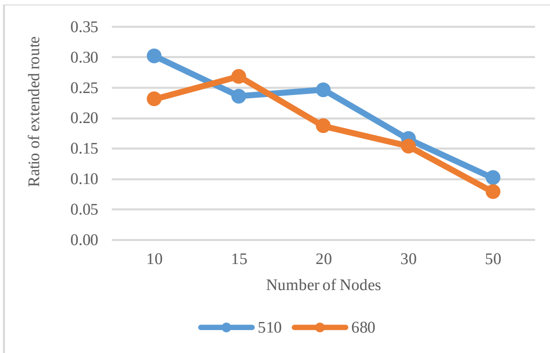

Figure 12. Ratio of extended routes in a 1250×1250 km² area space

Reviewing of Figure 11-a reveals that in the area space of 750×750 km2, increasing the number of nodes does not affect Nhub. However, by increasing the area space, the node number can affect the number of selected hubs. Moreover, Figure 11-b illustrates that a route number is related to the area space and node numbers.

Depending on the number of delivered parcels, nodes, and routes, the combination of vehicles was different in each case. As a general case, the number of cars on routes decreased in the bigger area, but the number of vehicles in the line hauls increased. Besides, the number of vehicles in TH=510 cases is more than that in TH=680 cases because TH=510 cases have fewer Nhub and Nroute. In Figure 13, the mean ratio of V-Line Haul to V-Route in different THs and area spaces is shown. It is apparent that this ratio is totally related to the area space and the value of TH, which is parallel with the results of Nhub in Figure 11-a. The number of hubs in TH=510 is less than that of TH=680, so the ratio of vehicles in the line hauls to the routes for TH=510 is less than that of TH=680.

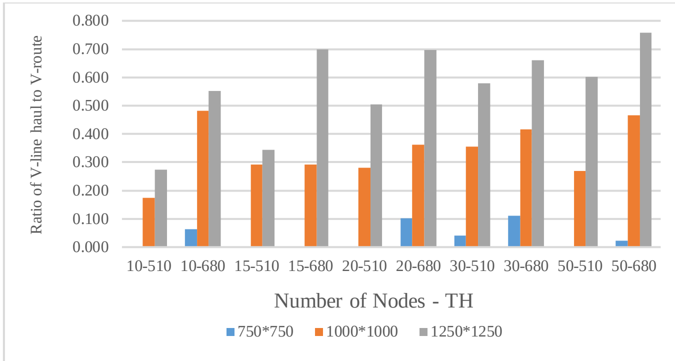

Figure 13. Ratio of V-Line Haul to V-Route in different TH and area space

Case study

In this section, two cases of road transportation in Iran are deliberated to validate the model's performance and method in actual cases. The first case comprises all 31 capital cities of Iran provinces, which was introduced by Bahrami et al. (2016). Adding 28 of Iran's most important or bordering cities, the second case with 59 nodes was designed and considered. According to business considerations, all data are standardized based on the real data of an Iranian door-to-door parcel delivery company. Both cases are solved five times in two-line haul distances of 510 and 680 km. The best results of each case are reported in Tables 6 and 7. It should be mentioned that bold digits demonstrate that the best final output of S2 is not yielded from its best solution of S1.

Table 6. Result of the solution method testing on two real cases - Part 1

Cases	TH	OFV			Nhub		Nroute		Ncon S1			Ncon S2		
		S1	S2	Diff	S1	S2	S1	S2	24	48	72	24	48	72
31	510	5212	4319	894	3	3	15	15	325	130	10	378	84	3
31	680	7971	**7948**	53	5	4	15	13	325	130	10	300	150	15
59	510	9445	**8416**	523	4	3	24	26	595	840	276	666	814	231
59	680	8546	8378	168	4	4	24	21	703	798	210	703	798	210

Table 7 Result of the solution method testing on two real cases - Prat 2

Cases	TH	Route S1	Line Haul S1	Route S2	Line Haul S2	Time

		L	M	S	L	M	S	L	M	S	L	M	S	S1	S2
31	510	21	4	8	12	2	2	21	5	5	10	6	0	374	331
31	680	15	4	8	12	6	10	12	4	6	16	2	6	357	334
59	510	15	6	11	12	16	0	17	10	12	10	4	2	1362	829
59	680	12	7	16	14	6	6	13	8	12	14	2	6	1246	878

The area space of the first case of 31 capital cities is more like the test instances with 1000×1000 km², so TH=510 provides better results. However, the area space of the second case is more similar to the instances with 1250×1250 km², so TH=680 offers better outcomes. The comparison of the results of S1 and S2 demonstrates that the method uses dissimilar efforts to cut the expenses. For example, it lessened the number of hubs in two cases, reduced the number of routes in three cases, increased the number of 24-hour transportations in two cases, and changed the combination of the vehicles in all cases.

Although allocated nodes or selected routes differ in each case, the number and location of hubs are more related to the maximum line hub distances. In both cases of 31 and 59 cities, the results TH=510 in the second stage (S2) indicate that three cities (i.e., Semnan, Qom, and Yazd) should be selected as the system hubs. The results of TH=680 in S2 suggest that four cities (i.e., Arak, Tehran, Shahrud, and Yazd) should be selected as the hubs. Hub location is a strategic decision that needs a medium or long-term investment in constructing hubs. The method can provide a powerful solution for the service provider since the results, especially the hub location, are unrelated to the number of nodes. If the company holds its strategy for the maximum distance of line hauls (TH), it can easily retain its hub location.

As mentioned before, the proposed solution method consists of two stages. To visually deliberate these stages, the final networks of the first and second stages of the case with 31 cities in TH=510 are illustrated in Figure 14 and Figure 15, respectively. Besides, the final network of the case with 59 cities in TH=680 is displayed in Figure 16. Since the performance of the first stage has a potential role in yielding better results in S2, in Figure 17 The convergence curve of the ABC algorithm for the second case with TH=680 is drawn as an example.

Comparing Figure 14 and Figure 15 demonstrates that the second stage has some beneficial effects on the final network. For example, Zahedan and Bushehr are allocated to Yazd, or the order of nodes in some routes is changed. Considering Figure 16, it is observed that the algorithm prolongs the distances of line hauls to avoid extended routes. Some interesting points of Figure 16 is the allocation of Hamedan and Sanandaj to the Tehran hub instead of the Arak hub (Arak is way closer to them), or the allocation of Taibad and Sarakhs to the Yazd hub instead of the Shahrud hub? The probable reasons can be their low demands or the space in their hubs' vehicles. Another point can be related to the allocation of Chabahar in a single-node route. Although Chabahar is close to Iranshahr, the method sends an empty vehicle instead of vehicles on the closer route. Since the demands of Chabahar are not high, the closer route vehicles may be full, or their travel time may exceed the extended time.

Comparing the best OFV of both cases demonstrates that covering some newly added cities in the second case causes the expenses of the system to be approximately double in comparison with the first case; however, the profit of running such a system is not considered here, and increasing the expenses will probably be reimbursed by new parcel deliveries. Besides, it is possible that covering such far and remote cities will not be profitable for the company.

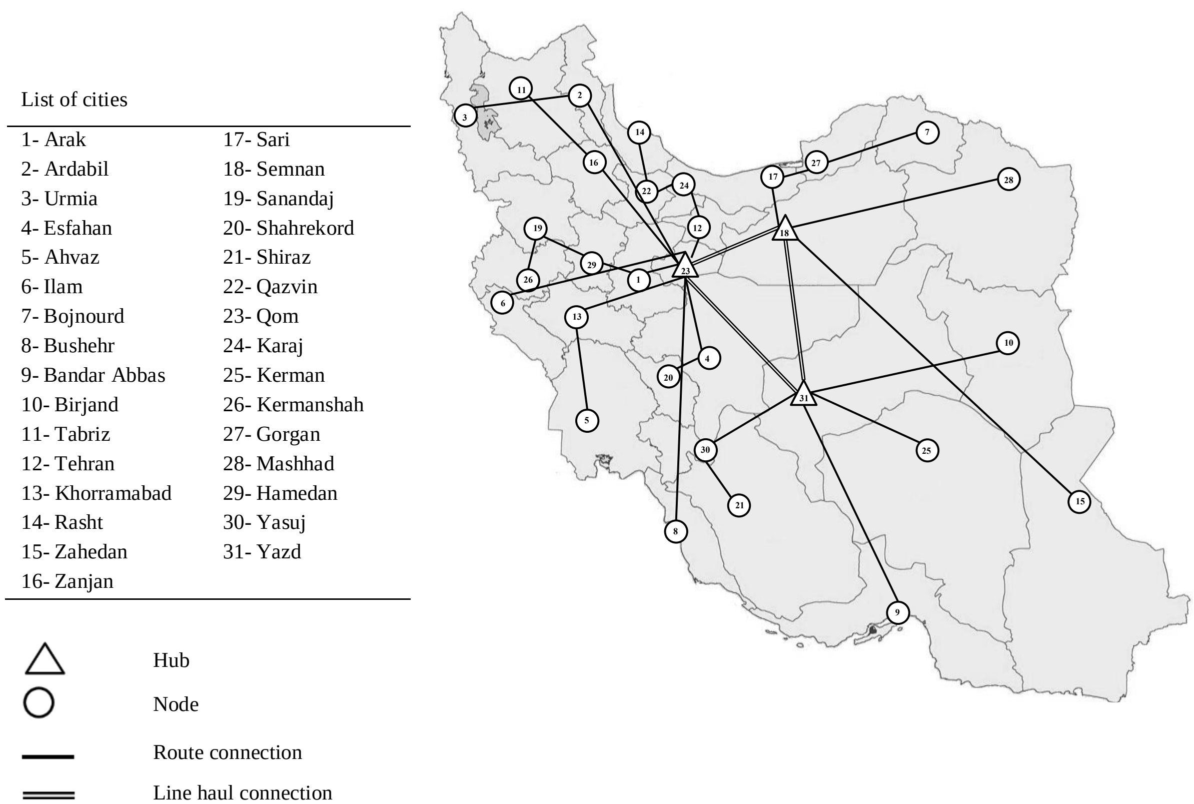

Figure 14. Final networks of the first stage of the case with 31 cities in TH=510

List of cities

1- Arak	17- Sari
2- Ardabil	18- Semnan
3- Urmia	19- Sanandaj
4- Esfahan	20- Shahrekord
5- Ahvaz	21- Shiraz
6- Ilam	22- Qazvin
7- Bojnourd	23- Qom
8- Bushehr	24- Karaj
9- Bandar Abbas	25- Kerman
10- Birjand	26- Kermanshah
11- Tabriz	27- Gorgan
12- Tehran	28- Mashhad
13- Khorramabad	29- Hamedan
14- Rasht	30- Yasuj
15- Zahedan	31- Yazd
16- Zanjan	

Hub

Node

Route connection

Line haul connection

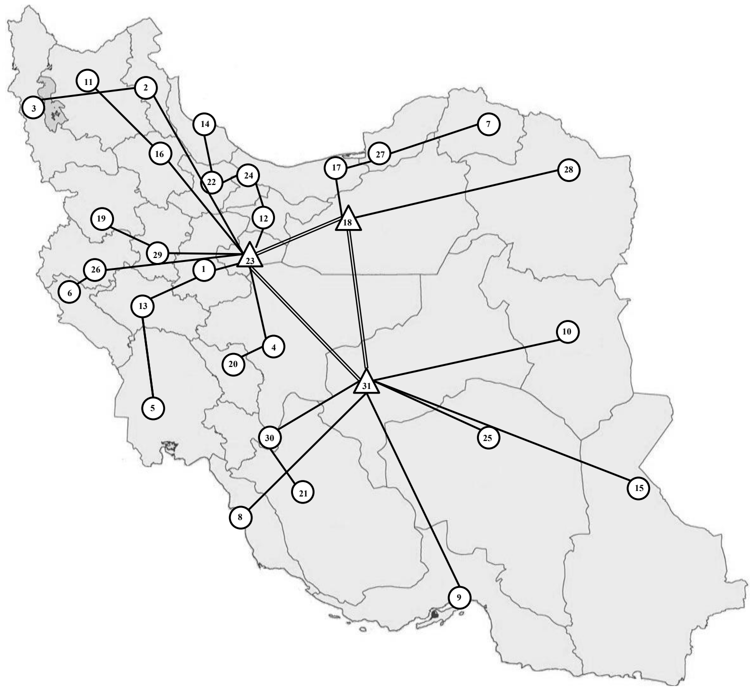

Figure 15. Final networks of the second stage of the case with 31 cities in TH=510

List of cities

1- Abadan	31- Khosravi
2- Astara	32- Dashli borun
3- Arak	33- Ramsar
4- Ardabil	34- Rasht
5- Urmia	35- Zahedan
6- Esfahan	36- Zanjan
7- Aslan Duz	37- Sari
8- Ahvaz	38- Sarakhs
9- Iranshahr	39- Sarv
10- Ilam	40- Semnan
11- Babol	41- Sanandaj
12- Babolsar	42- Shahrud
13- Bajgiran	43- Shahrekord
14- Bazargan	44- Shiraz
15- Bojnord	45- Qazvin
16- Bandar Emam	46- Qom
17- Bandar Anzali	47- Kashan
18- Boshehr	48- Karaj
19- Bandar Abbas	49- Kerman
20- Birjand	50- Kermanshah
21- Bile Savar	51- Gorgan
22- Piranshahr	52- Lar
23- Taibad	53- Mashhad
24- Tabriz	54- Mahabad
25- Tehran	55- Mirjaveh
26- Jolfa	56- Neyshabur
27- Chabahar	57- Hamedan
28- Chalus	58- Yasuj
29- Khorramabad	59- Yazd
30- Khorramshahr	

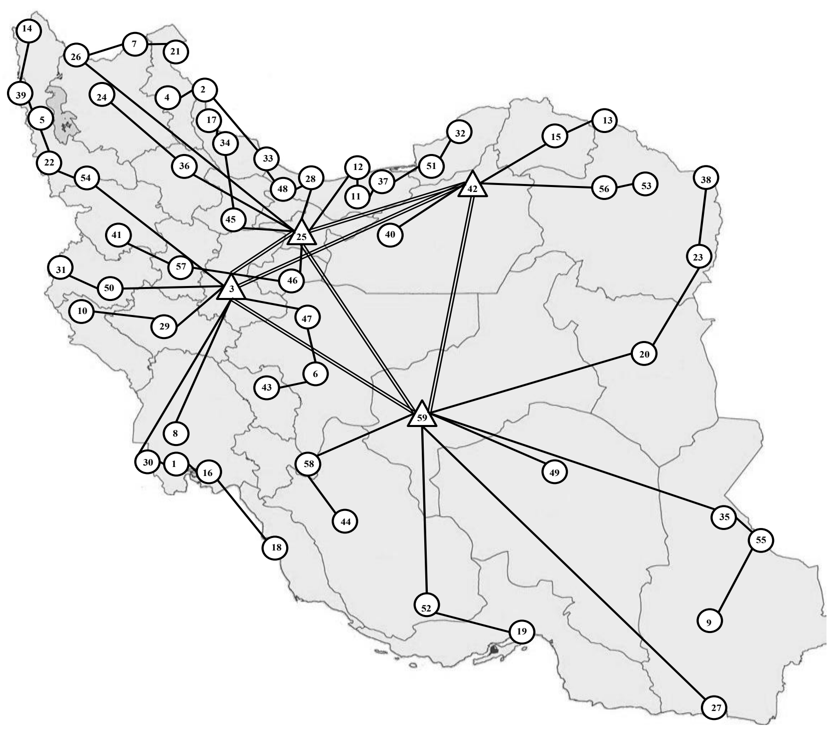

Figure 16. Final networks of the second stage of the case with 59 cities in TH=680

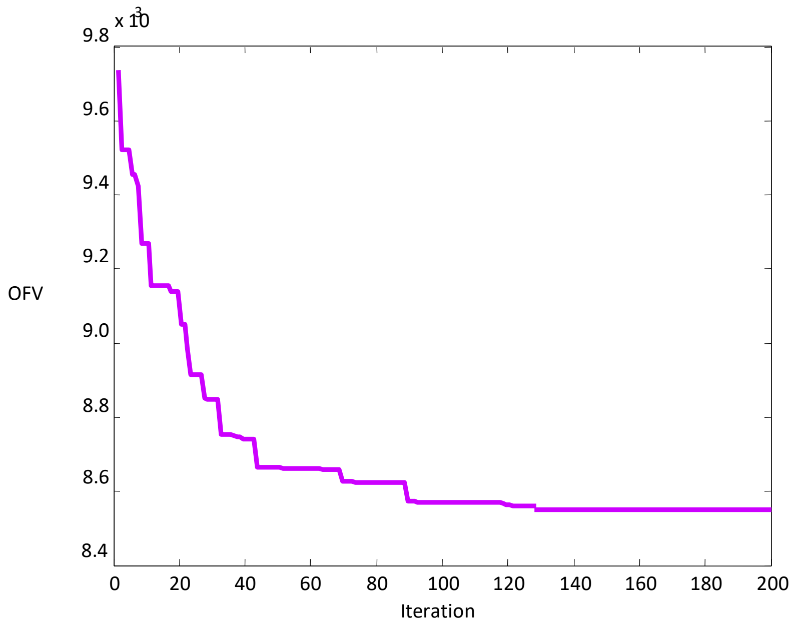

Figure 17. Convergence curve of the ABC for the second case with TH=680

Conclusion

In this paper, a variant of the many-to-many location-routing problem (MMLRP) was considered to find the number and location of hubs, allocation of nodes to hubs, path routing, and the number of vehicles in routes and line hauls. Inspired by the real cases of Iran's door-to-door delivery services, multi-path routing, heterogeneous cars, and time windows were considered in the MMLRP. A novel polynomial-sized mixed-integer programming model that minimized the cost of establishing a logistic provider in a broad and sparse country with restricted and incomplete roads and different vehicle capacity sizes was presented. As logistic service providers in the parcel services confront a more complex situation than other services, a new two-stage hybrid method based on two meta-heuristic algorithms, namely artificial bee colony (ABC) and simulated annealing (SA), was presented to solve the new NP-hard model. Assessing the proposed model and the solution method, 75 test instances were generated based on the practical assumptions with different numbers of 10 to 50 nodes in three area spaces of 750×750, 1000×1000, and 1250×1250 km². The computational results indicated the high performance of the solution method compared to CPLEX and an SA-based method. A sensitive analysis of test instances in two-line haul distances was proposed to provide more detail about the embedded parameters. It verified the effects of the second stage, which decreased the OFV by using different devices, such as reducing the number of hubs and routes, changing the combination of vehicle fleets, or increasing 24-hour transportation. It also revealed that the area space has an unavoidable impact on all model parameters.

Furthermore, two real cases, with 31 capital cities of Iran provinces and 59 important and bordering cities of Iran, were considered based on data from a parcel delivery company in Iran. The results were illustrated on the Iran map to help the proposed problem be observed better visually. The outcomes proved that the first case is more similar to test instances with the area space of 1000×1000 km2, while the second case is like those with 1250×1250 km2. It was shown that if the company owners strategically decided on the line haul distances, covering new cities in the future, they may only redesign the routing regime, allocate new cities to the current hubs, or change the combination of vehicle fleets. Still, it will not extensively affect the hub number or locations. Therefore, service providers can easily expand their parcel delivery network without substantial augmentation of capital investment in hub locations. Consequently, when considering only capital cities, it is better to set the maximum distance of line hauls to 510 km with three hubs. At the same time, if the company decides to expand its network in the future, it is better to increase the maximum distances of line hauls to 680 km and establish four hubs.

Data Availability Statement

Data available on request from the author.

Acknowledgements

The author would like to thank all the participants in the present study.

Ethical considerations

The author have witnessed the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancy.

Funding

The author received no financial support for this article's research, authorship, and/or publication.

Conflict of interest

The author declares no potential conflict of interest regarding the publication of this work.

References

Abbasi, M., Mokhtari, N., Shahvar, H., & Mahmoudi, A. (2019). Application of variable neighborhood search for solving large-scale many-to-many hub location routing problems. *Journal of Advances in Management Research*, *16*(5), 683–697.

Aloullal, A., Saldanha-da-Gama, F., & Todosijević, R. (2023). Multi-period single-allocation hub location-routing: Models and heuristic solutions. *European Journal of Operational Research*, *310*(1), 53–70.

Andaryan, A. Z., Mousighichi, K., & Ghaffarinasab, N. (2024). A heuristic approach to the stochastic capacitated single allocation hub location problem with Bernoulli demands. *European Journal of Operational Research*, *312*(3), 954–968.

Athamneh, R. A., Tanash, M., Hani, D. B., Rawshdeh, M., Alawin, A., & Albataineh, Z. (2023). Variable Neighborhood Search Algorithm for the Single Assignment Incomplete Hub Location Problem with Modular Capacities and Direct Connections. *Operations Research Perspectives*, *11*, 100286.

Bahrami, F., Safari, H., Tavakkoli-Moghaddam, R., & Modarres Yazdi, M. (2016). On modeling door-to-door parcel delivery services in Iran. *Interdisciplinary Journal of Management Studies*, *9*(4), 883–906.

Bahrami, F., Safari, H., Tavakkoli-Moghaddam, R., & Modarres Yazdi, M. (2017). Road hub location-routing issue in a sparse and distant area. *Industrial Management Journal*, *9*(1), 59–78.

Basirati, M., Akbari Jokar, M. R., & Hassannayebi, E. (2020). Bi-objective optimization approaches to many-to-many hub location routing with distance balancing and hard time window. *Neural Computing and Applications*, *32*(17), 13267–13288.

Bement, T. R. (1989). Taguchi techniques for quality engineering. In: Taylor & Francis.

Brimberg, J., Todosijević, R., Urošević, D., & Mladenović, N. (2021). Efficient flow models for the uncapacitated multiple allocation p-hub median problem on non-triangular networks. *Computers & Industrial Engineering*, *162*, 107723.

Bütün, C., Petrovic, S., & Muyldermans, L. (2021). The capacitated directed cycle hub location and routing problem under congestion. *European Journal of Operational Research*, *292*(2), 714-734.

Černý, V. (1985). Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of optimization theory and applications*, *45*(1), 41-51.

Čupić, A., & Teodorović, D. (2014). A multi-objective approach to the parcel express service delivery problem. *Journal of Advanced Transportation*, *48*(7), 701-720.

Davari, S., Zarandi, M. F., & Turksen, I. (2013). The incomplete hub-covering location problem considering imprecise location of demands. *Scientia Iranica*, *20*(3), 983-991.

De Camargo, R. S., de Miranda Jr, G., O'Kelly, M. E., & Campbell, J. F. (2017). Formulations and decomposition methods for the incomplete hub location network design problem with and without hop-constraints. *Applied Mathematical Modelling*, *51*, 274–301.

De Sá, E. M., Morabito, R., & de Camargo, R. S. (2018). Benders decomposition applied to a robust multiple allocation incomplete hub location problem. *Computers & Operations Research*, *89*, 31–50.

Estrada-Romeu, M., & Robusté, F. (2015). Stopover and hub-and-spoke shipment strategies in less-than-truckload carriers. *Transportation Research Part E: Logistics and Transportation Review*, *76*, 108–121.

Fallah-Tafti, M., Honarvar, M., Tavakkoli-Moghaddam, R., & Sadegheih, A. (2022). Mathematical modeling of a bi-objective hub location-routing problem for rapid transit networks. *RAIRO-Operations Research*, *56*(5), 3733–3763.

Ghodratnama, A., Tavakkoli-Moghaddam, R., & Azaron, A. (2013). A fuzzy possibilistic bi-objective hub covering problem considering production facilities, time horizons, and transporter vehicles. *The International Journal of Advanced Manufacturing Technology*, *66*(1), 187-206.

Gianpaolo, G., Gilbert, L., & Roberto, M. (2013). Introduction to Logistics Systems Management. In: Wiley.

Gu, W., Archetti, C., Cattaruzza, D., Ogier, M., Semet, F., & Speranza, M. G. (2024). Vehicle routing problems with multiple commodities: A survey. *European Journal of Operational Research*, *317*(1), 1–15.

Guo, F., Wang, Z., Huang, Z., & Ma, X. (2024). Robust optimization of microhub network and mixed service strategy for a multidepot location-routing problem. *Computers & Industrial Engineering*, *190*, 110070.

Kabadurmus, O., & Erdogan, M. S. (2023). A green vehicle routing problem with multi-depot, multi-tour, heterogeneous fleet and split deliveries: a mathematical model and heuristic approach. *Journal of Combinatorial Optimization*, *45*(3), 89.

Karaboga, D., & Akay, B. (2009). A comparative study of artificial bee colony algorithm. *Applied mathematics and computation*, *214*(1), 108–132.

Karaoglan, I., Altiparmak, F., Kara, I., & Dengiz, B. (2012). The location-routing problem with simultaneous pickup and delivery: Formulations and a heuristic approach. *Omega*, *40*(4), 465-477.

Kartal, Z., Hasgul, S., & Ernst, A. T. (2017). Single allocation p-hub median location and routing problem with simultaneous pick-up and delivery. *Transportation Research Part E: Logistics and Transportation Review*, *108*, 141–159.

Khoshgehbari, F., & Al-e, S. M. J. M. (2023). Ambulance location routing problem considering all sources of uncertainty: Progressive estimating algorithm. *Computers & Operations Research*, *160*, 106400.

Kirkpatrick, S., Gelatt Jr, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *science*, *220*(4598), 671-680.

Lüer-Villagra, A., Eiselt, H. A., & Marianov, V. (2019). A single allocation p-hub median problem with general piecewise-linear costs in arcs. *Computers & Industrial Engineering*, *128*, 477–491.

Nagy, G., & Salhi, S. (1998). The many-to-many location-routing problem. *Top*, *6*(2), 261-275.

Nielsen, P., Dahanayaka, M., Perera, H. N., Thibbotuwawa, A., & Kilic, D. K. (2024). A systematic review of vehicle routing problems and models in multi-echelon distribution networks. *Supply Chain Analytics*, *7*, 100072.

O'Kelly, M. E. (1986). The location of interacting hub facilities. *Transportation science*, *20*(2), 92–106.

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

O’Kelly, M., Sun, X., & Wandelt, S. (2025). Hub location problems: A meta review and ten disruptive research challenges. *Journal of the Air Transport Research Society*, 100073.

Oudani, M., Aouam, T., & Ghogho, M. (2023). The p-hub centre routing problem with emissions budget: Formulation and solution procedure. *Computers & Operations Research*, *154*, 106211.

Öztürk, C., Tuzkaya, G., & Bulkan, S. (2021). Centrality based solution approaches for median-type incomplete hub location problems. *Computers & Industrial Engineering*, *156*, 107275.

Pourmohammadi, P., Tavakkoli-Moghaddam, R., Rahimi, Y., & Triki, C. (2023). Solving a hub location-routing problem with a queue system under social responsibility by a fuzzy meta-heuristic algorithm. *Annals of Operations Research*, *324*(1), 1099–1128.

Ratli, M., Urošević, D., El Cadi, A. A., Brimberg, J., Mladenović, N., & Todosijević, R. (2022). An efficient heuristic for a hub location routing problem. *Optimization Letters*, *16*(1), 281-300.

Rieck, J., Ehrenberg, C., & Zimmermann, J. (2014). Many-to-many location-routing with inter-hub transport and multi-commodity pickup-and-delivery. *European Journal of Operational Research*, *236*(3), 863–878.

Shang, C., Ma, L., & Liu, Y. (2023). Green location routing problem with flexible multi-compartment for source-separated waste: A Q-learning and multi-strategy-based hyper-heuristic algorithm. *Engineering Applications of Artificial Intelligence*, *121*, 105954.

Sharma, A., Jakhar, S. K., Vlachos, I., & Kumar, S. (2025). Advances in hub location problems: a literature review and research agenda. *International Journal of Productivity and Performance Management*, *74*(1), 24–55.

Shi, Y., Lin, Y., Wang, S., Wen, H., Lim, M. K., & Li, Y. (2023). A simultaneous facility location and vehicle routing problem with recyclable express packaging consideration for sustainable city logistics. *Sustainable Cities and Society*, *98*, 104857.

Vahdani, B., Tavakkoli-Moghaddam, R., Zandieh, M., & Razmi, J. (2012). Vehicle routing scheduling using an enhanced hybrid optimization approach. *Journal of Intelligent Manufacturing*, *23*(3), 759-774.

Wandelt, S., Wang, S., & Sun, X. (2025). A literature review on hub location-routing models and their solution techniques. *Computers & Operations Research*, *173*, 106861.

Wang, C., Liu, Y., & Yang, G. (2023). Adaptive distributionally robust hub location and routing problem with a third-party logistics strategy. *Socio-Economic Planning Sciences*, *87*, 101563.

Wang, S., Sun, W., Baldacci, R., & Elomri, A. (2025). Exact solution of location–routing problems with heterogeneous fleet and weight-based carbon emissions. *Transportation Research Part E: Logistics and Transportation Review*, *193*, 103862.

Wang, X., & Li, X. (2017). Carbon reduction in the location routing problem with heterogeneous fleet, simultaneous pickup-delivery, and time windows. *Procedia Computer Science*, *112*, 1131-1140.

Wasner, M., & Zäpfel, G. (2004). An integrated multi-depot hub-location vehicle routing model for network planning of parcel service. *International Journal of Production Economics*, *90*(3), 403–419.

Wu, Y., Qureshi, A. G., & Yamada, T. (2022). Adaptive large neighborhood decomposition search algorithm for multi-allocation hub location routing problem. *European Journal of Operational Research*, *302*(3), 1113–1127.

Yaman, H., Kara, B. Y., & Tansel, B. Ç. (2007). The latest arrival hub location problem for cargo delivery systems with stopovers. *Transportation Research Part B: Methodological*, *41*(8), 906-919.

Zhao, Q., Wang, W., & De Souza, R. (2018). A heterogeneous fleet two-echelon capacitated location-routing model for joint delivery arising in city logistics. *International Journal of Production Research*, *56*(15), 5062–5080.

Zhao, Z., Lee, C. K., Tsang, Y. P., & Xu, X. (2025). A heuristic-attention method for location-routing problems with shared pick-up stations in green last-mile delivery. *Transportation Research Part C: Emerging Technologies*, *172*, 105031.