Computers & Operations Research 152 (2023) 106126

Contents lists available at ScienceDirect

Computers and Operations Research

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

Home healthcare staff dimensioning problem for temporary caregivers: A matheuristic solution approach

Erfaneh Nikzad [a], Mahdi Bashiri [b,*], Babak Abbasi [c]

[a] Industrial Engineering Department, Faculty of Engineering, Shahed University, Tehran, Iran
[b] Centre for Business in Society, Coventry University, Coventry, United Kingdom
[c] School of Accounting, Information Systems and Supply Chain, RMIT University, Melbourne, Australia

ARTICLE INFO

Keywords:
Home health care
Staff dimensioning
Matheuristic algorithm
Two-stage stochastic programming

ABSTRACT

Staff dimensioning, defined as determining the required numbers of caregivers with different types of skills, is a key decision for home healthcare systems. Home healthcare providers often use a combination of permanent and temporary (casual) caregivers. Determining the required number of temporary caregivers with different skill sets considering uncertainty and routing cost is the main objective of this study. To this end, we propose a two-stage stochastic programming model for the staff dimensioning problem for temporary caregivers, taking into account uncertainties in the required class of service, the required number of visits, and the required service time for each patient. Staff dimensioning decisions are defined in the first stage, and assignment with routing are positioned in the second stage of the model. To solve the problem, a two-phase matheuristic algorithm is developed where an initial solution is generated in the first phase by using an intermediate mathematical model and solving a series of Traveling Salesman Problems (TSPs), then a fix-and-optimize strategy is developed in the second phase to improve the obtained solution. The efficiency of the proposed matheuristic algorithm is examined by various test problems. The results highlight that the proposed model and solution method can be used by HHC providers to effectively utilize the option of recruitment of temporary caregivers in their resource planning considering inevitable uncertain parameters.

1. Introduction

Home healthcare (HHC) systems provide medical and paramedical services by sending caregivers to patients' homes. Providing care at home increases patient satisfaction significantly and improves the quality of the delivered services. Moreover, many countries face aging populations and increasing healthcare costs to meet their needs. Therefore, HHC services are growing rapidly worldwide, especially in developed countries such as France, UK, and USA, to reduce healthcare costs and to improve the quality of healthcare services (Yuan et al., 2018).

Staff dimensioning is an important issue in healthcare systems (Andersen et al., 2019; Vieira et al., 2018). The facts that patients in HHC must be visited in specified time windows and the caregivers require to travel between the patients' homes further complicate the staff dimensioning compare to others.

Staff dimensioning is a long-term decision and in HHC systems, caregivers with different types of skills are needed due to the wide variety of required care services. Compatibility between patients' required services and caregivers' skills needs to be considered in the HHC. In some cases, because of the uncertainty of some parameters such as travel and service times, etc, the number of available caregivers is not enough to visit all patients. Therefore, in these cases, a provider of HHC services can hire temporary caregivers who can be the trained staff from other departments to be used for HHC services. In this study, two types of caregivers including permanent and temporary caregivers are considered. The number of available permanent caregivers is considered a predefined input parameter in this study. While temporary caregivers are used to deal with uncertainty, so their hiring are considered a short or medium term decision. The required number of temporary caregivers in each week is determined at the beginning of the week. So, the number of temporary caregivers should be determined in this study.

Continuity of care, defined as visiting a patient by the same caregiver, increases both patient and caregiver satisfaction. Continuity of care has been considered as either a soft or hard constraint in some previous studies. Furthermore, balancing the number of patients assigned to each caregiver can also increase the caregiver's satisfaction. These features are considered in the proposed model, to take into

* Corresponding author.
E-mail addresses: e.nikzad@shahed.ac.ir (E. Nikzad), Mahdi.Bashiri@coventry.ac.uk (M. Bashiri), babak.abbasi@rmit.edu.au (B. Abbasi).

https://doi.org/10.1016/j.cor.2022.106126
Received 18 April 2020; Received in revised form 1 June 2022; Accepted 15 December 2022
Available online 11 January 2023

account patient's and caregiver's satisfaction as well as the cost-related concerns. Moreover, considering the home locations of caregivers as the start and end points of their corresponding routes is needed, to reflect the routing cost and time.

As mentioned before, patients in HHC systems may require different types of services. Patients who require same service can be categorized into different groups (or classes) according to their required service level. The required type of service for each patient is known but her/his class is determined by scheduling an appointment or receiving a call before starting service to determine the service level required by patients.[1] Therefore, the class of each patient is not known in advance but can be determined through several scenarios. The required number of visits for each patient depends on their required services and their class, so this parameter is also considered to be uncertain. In addition, the service time is considered as another uncertain parameter that is realized in various scenarios.

The contributions of this study can be summarized as follows. A new two-stage stochastic programming model for staff dimensioning is proposed to deal with uncertainties. In the proposed model, decisions regarding staff dimensioning, scheduling, assignment, and routing are considered, simultaneously. A matheuristic algorithm comprising two phases is proposed to solve the model, efficiently. A near-optimal solution is extracted in the first phase by using an iterative procedure hybridized with an intermediate integer programming model and a series of TSPs. In the second phase, the obtained solutions corresponding to scheduling, assignment and routing decisions are improved with a fix-and-optimize procedure.

This paper is organized as follows. Studies related to resource planning in HHC are briefly reviewed in Section 2. The proposed mathematical formulation is presented in Section 3, and the developed matheuristic algorithm is explained in Section 4. Section 5 is devoted to numerical experiments. Finally, conclusions and perspectives are provided in the last section.

2. Literature review

Resource planning in HHC comprises districting, staff dimensioning, assignment, and routing decisions. Most studies in this context are related to assignment and routing problems (e.g. Cappanera et al., 2018; Grenouilleau et al., 2019; Decerle et al., 2018; Moussavi et al., 2019). The studies on scheduling and routing problems in HHC were reviewed by Fikar and Hirsch (2017) and Cissé et al. (2017). Despite its importance the resource planning aspect of HHC as received limited attention (Nikzad et al., 2020). However, integrating decisions related to resource planning in home healthcare helps HHC services to be more efficient and less costly. Recent studies related to resource planning in HHC are summarized in Table 1 and compared with this study.

As shown in Table 1, most studies related to resource planning in HHC focused on assignment and routing decisions, while resource planning decisions have been rarely considered simultaneous in the previous studies. However, comprehensively considering these decisions may decrease the total cost of HHC systems. Therefore, a mathematical model is proposed to consider staff dimensioning, scheduling, assignment, and routing decisions, simultaneously.

Table 1 further highlights that a variety of uncertainties, such as travel and service times, demand, availability of caregivers, and required types of skills, have been considered in the studies on resource planning in HHC. Dealing with uncertainties is a great challenge for decision makers. Stochastic and robust programming are two well-known frameworks to consider uncertainty. Stochastic programming is used when the probability distribution corresponding to an uncertain parameter is known. Its two main categories are two-stage or multi-stage programming, and chance constraint programming (Birge and Louveaux, 2011). Strategic decisions are determined more accurately by implementing the two-stage stochastic programming framework because it considers the expected value of the costs corresponding to recourse actions.

As mentioned in Table 1, stochastic and robust programming have been widely used for resource planning in HHC. Lanzarone and Matta (2014), Carello and Lanzarone (2014), and Cappanera et al. (2018) proposed robust optimization for assignment and routing problems in HHC by considering demand uncertainty. Nguyen et al. (2015) proposed a robust model for assignment and routing problems by considering the availability of caregivers as an uncertain parameter. Two-stage stochastic programming is another approach used to deal with uncertainties when the probability distributions of uncertain parameters can be obtained from historical data (e.g. Yuan et al., 2015, 2018; Shi et al., 2018). In this study, a two-stage stochastic model is developed to consider the uncertainty of classes (or type) of patients, the required number of visits, and service time. The required number of temporary caregivers is the first stage decision variable. The scheduling, assignment, and routing decisions are considered as the second stage decision variables.

Table 1 also shows that matheuristic algorithms have been developed widely in the studies related to resource planning in HHC. Grenouilleau et al. (2019) proposed a matheuristic algorithm for routing and scheduling problems in HHC based on a combination of set-partitioning model and Large Neighborhood Search (LNS) algorithm, where the routes which are used in the set-partitioning problem are generated by the LNS algorithm. Moussavi et al. (2019) developed a matheuristic algorithm for assignment and routing problems in HHC. Their algorithm contains three steps in which the number of caregivers for each day is determined in the first phase and the set of patients who must be served each day is obtained in the second phase. Finally, the assignment of patients to caregivers and routing decisions are determined in the last phase. Allaoua et al. (2013) designed a two-phase matheuristic algorithm for routing and staff rostering problems. The feasible routes are generated in the first phase which are used in the set-partitioning-based model in the second phase to make staff rostering decisions. Fikar and Hirsch (2015) developed a two-phase matheuristic algorithm. First, feasible walking routes are determined by set-partitioning and then an extended biased randomized savings heuristic and the Tabu search determine the schedule, vehicle routes and improve walking routes. Nguyen et al. (2015) proposed a hybrid algorithm based on a mathematical model and genetic algorithm (GA) for robust assignment routing problems which considers the availability of caregivers as an uncertain parameter. Also, Cappanera et al. (2018) proposed a matheuristic algorithm for solving robust assignment and routing problems in HHC. It can be concluded that most matheuristic algorithms for solving assignment and routing problems in HHC are designed by generating a set of feasible routes and then selecting routes by optimizing a set-partitioning model. In our proposed model, a routing decision is considered as the second-stage variable that depends on scenarios. So, by increasing the number of scenarios, the number of feasible routes will increase, significantly. Therefore, the solution approaches proposed in previous studies seem to be inefficient to solve the proposed model. In this study, a two-phase matheuristic algorithm is developed. An initial solution is generated in the first phase and the scheduling, assignment, and routing decisions are improved in the second phase.

Matheuristic algorithms are widely used in the vehicle routing problems (VRP) related research such as production routing problems (PRP), inventory routing problems (IRP) and etc. (e.g. Bertazzi et al., 2019; Miranda et al., 2018; Li et al., 2019; Hernandez et al., 2019; Yu et al., 2019; Solyalı and Süral, 2017). Recently, matheuristic algorithms have been developed based on fix-and-optimize procedures. In these algorithms, the values of some decision variables are dictated and other

[1] https://www.medicare.gov/what-medicare-covers/whats-home-health-care

Table 1
Studies related to resource planning in HHC. The abbreviations are: D: Districting, S: Staff dimensioning, A: Assignment, R: Routing, MP: Multi-period, CC: Continuity of care, C: Compatibility between caregivers' skills and patients required services, NP: Number of patients assigned to each caregiver, PR: Patients preferences on caregivers assignment, VR: Visiting regulations, math: Matheuristic algorithm, h: Heuristic algorithm, BP: Branch and price, BC: Branch and cut metah: Meta-heuristic, SR: Skill requirements, CP: Classes of patients based on their symptoms, RV: Required number of visits, ST: Service time, TT: Travel time, LSA: L-shaped algorithm, ALNS: Adaptive large neighborhood search.

References	Types of Decisions				Types of uncertainties		Uncertain parameters	Assumptions						Solution method
	D	S	A	R	Stochastic	Robust		MP	CC	C	NP	PR	VR	
Grenouilleau et al. (2019)			×	×				×	×	×		×		math
Cappanera et al. (2018)			×	×		×	Demand	×	×	×				math
Yuan et al. (2018)				×	×		TT & ST							BP
Decerle et al. (2018)				×						×				metah
Lin et al. (2018)			×	×				×		×				metah
Shi et al. (2018)				×	×		TT & ST							metah
Braekers et al. (2016)			×	×						×				metah
Yuan et al. (2015)				×	×		ST & SR							BP
Rodriguez et al. (2015)		×	×	×	×		Demand			×				math
Yalçındağ et al. (2016)			×	×				×	×	×				math
Bahadori-Chinibelagh et al. (2019)			×	×										h
Zhan et al. (2020)			×	×	×		ST							math & LSA
Hashemi Doulabi et al. (2020)			×	×	×		TT & ST			×				LSA & BC
Fathollahi-Fard et al. (2020a)			×	×										metah
Fathollahi-Fard et al. (2020b)			×	×		×	TT & ST	×						metah
Cinar et al. (2021)			×				×							math & ALNS
Nikzad et al. (2020)	×	×	×	×	×		ST & TT	×	×	×				math
This study		×	×	×	×		CP & RV & ST	×	×	×	×	×	×	math

decision variables are optimized to improve their values. Neves-Moreira et al. (2019) and Li et al. (2019) designed matheuristic algorithms based on a fix-and-optimize procedure in the PRP. Campelo et al. (2019) developed a matheuristic algorithm based on this procedure for a consistent vehicle routing problem. Also, Lindahl et al. (2018) and Dorneles et al. (2014) developed a matheuristic algorithm for timetabling problems based on a fix-and-optimize procedure.

As mentioned, staff planning in HHC services is a crucial decision that significantly affects on the cost of services, as well as patients' and caregivers' satisfaction. To the best of the author's knowledge, Rodriguez et al. (2015) and Nikzad et al. (2020) are the only studies related to staff dimensioning in HHC. Rodriguez et al. (2015) proposed a two-stage mixed-integer programming model that minimizes the required number of caregivers in a single period when the demand for each type of activity is uncertain. Rodriguez et al. (2015) developed a two-phase algorithm for solving their proposed model in which the minimum required number of caregivers with different types of skills is determined under each scenario. Then the optimal number of caregivers is obtained in the second phase. Rodriguez et al. (2015) determined the number of caregivers in each profession required to obtain a certain performance level by considering total demand for different activities as an uncertain parameter. Routing costs between sectors and patients in the same sector are not mentioned in their study, while in our study the required number of caregivers in each profession is determined based on hiring and routing costs when total demand is served. Also, the class of a patient's required services, the required number of visits, and the service time are assumed to be as uncertain parameters in our study. The continuity of care and regulations for visiting patients, two crucial aspects of HHC planning, were not considered by Rodriguez et al. (2015). These aspects can significantly affect the required number of caregivers at each level of skill, so they are considered in this study. Also, patient preferences on caregivers assignment are considered in this study, which was also not mentioned in the model proposed by Rodriguez et al. (2015).

Our work is closely related to that of Nikzad et al. (2020), who proposed a two-stage stochastic programming model for districting in HHC which considered staff dimensioning, scheduling, and routing decisions, simultaneously. They considered service and travel times as uncertain parameters. However, unlike this paper, patient preferences on caregiver assignment and regulations for visiting patients (according to medical prescription) were not considered in their study. They proposed a four-phase matheuristic algorithm. In the first phase, a size-reduction approach was used to eliminate the set of the potential number of districts and visiting sequences. Then a cluster-based mathematical model was used to generate an initial solution. They used the Progressive Hedging combined with the Frank–Wolfe algorithm to solve the cluster-based model. Then the unfeasible solutions obtained for districting were repaired with a fix-and-optimize procedure. Routing and assignment decisions were improved in the last phase.

3. Problem description and formulation

The staff dimensioning problem in HHC services is to determine the required number of caregivers with different types of skills to meet patients' demands. Staff dimensioning is a strategic decision and is determined without revealing uncertain parameters. Because of facing with uncertainty, sometimes the number of hired caregivers is not enough to meet all patients' demands. In these cases, HHC centers can hire additional caregivers for a short time such as a week, or assign some trained staff from other departments to provide HHC services, temporarily to deal with uncertainty. Therefore, in this paper, two types of caregivers including permanent and temporary are considered and the required number of temporary caregivers with different types of skills during the planning horizon (one week) is determined based on hiring and routing costs as well as some factors which affect patients and caregivers satisfaction considering several uncertain parameters.

The staff dimensioning problem is defined by a graph $G(\mathcal{V}, \mathcal{L})$, where $\mathcal{V} = \{1, 2, \dots, N, N+1, N+2, \dots, N+K\}$ is the set of nodes and $\mathcal{L} = \{(i,j) | i, j \in \mathcal{V}, i \neq j\}$ is the set of arcs. The nodes in set $\mathcal{N} = \{1, 2, \dots, N\}$ denote patients, and the nodes in set $\mathcal{N}_c = \{N+1, N+2, \dots, N+K\}$ denote the locations of the caregivers, where $\mathcal{K} = \{1, 2, \dots, K\}$ is the set of caregivers and $\mathcal{K}_P = \{1, 2, \dots, K_1\}$ and $\mathcal{K}_T = \{K_1+1, K_1+2, \dots, K\}$ are sets of permanent and temporary caregivers, respectively. Also, node $N+k$ determines the origin and destination of the route for the kth caregiver.

HHC centers provide different types of services and they schedule an appointment or call before starting service to determine care and services required by patients.[2] Therefore, patients require different types of services defined by the set $\mathcal{F} = \{1, 2, \dots, F\}$. Different classes can be considered for patients based on their required care services. The set of different classes for each type of service is defined by set $\mathcal{A} = \{1, 2,$

[2] https://www.medicare.gov/what-medicare-covers/whats-home-health-care

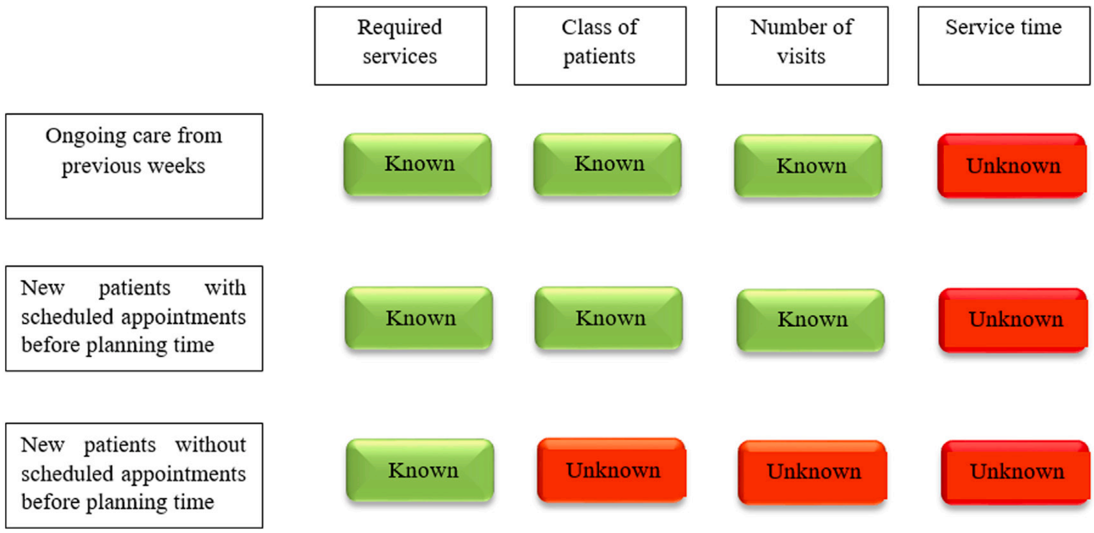


Fig. 1. Different categories of patients and available information for each category.

$\ldots, A\}$. The patients available in HHC centers at the beginning of each week are classified into three categories. The first category consists of patients whose care is continued from previous weeks, therefore, the required services and their level (classes of patients) are known. New patients with scheduled appointments (known classes of patients) are in the second category. The third category consists of new patients that the preliminary appointment has not been done, so there is no enough information about their required level of service at the beginning of the planning horizon. Known and unknown information about each category of patients is illustrated in Fig. 1. As illustrated in this figure, the class of patients is unknown in advance (before the appointment) for some of the patients, so they are considered as uncertain parameters. The required number of visits for each patient depends on the class of the patient and considered as another uncertain parameter. Service time $st_i(\xi)$ is considered as another uncertain parameter in this study. The parameters which have been considered in the current research can be categorized as parameters related to the caregivers, patients, and features that are described as follows.

- **Caregivers**. Caregivers with different skills, such as physicians and nurses, provide services to patients at home. The compatibility between their skills and patients' required services is considered according to their class. Each caregiver $k \in \mathcal{K}$ is defined by hiring cost h_k and his/her skill level which is defined by a binary parameter q_{kfa} that is equal to 1 if the kth caregiver has proper skill level to meet the patients that require service type f in the ath class. Also, a maximum working time ϕ is considered for the caregivers.
- **Patients**. Each patient $i \in \mathcal{N}$ is defined by his/her required service $r_{ifa}(\xi)$ that is a binary parameter that is equal to 1 if the ith patient who is in the ath class requires service type f under scenario ξ. Each patient has a number of required visits $\lambda_i(\xi)$ that should be scheduled during the planning horizon. A minimum required duration between two consecutive visits δ_{fa} is considered for scheduling the visits. b_i is the latest allowable service starting time of the ith patient.
- **Features**. Some aspects are considered to increase patients' satisfaction such as continuity of care and considering patients' preferences for assigning caregivers. Continuity of care is considered as a soft constraint and cost θ_i should be paid per caregiver who is assigned to the ith patient. For each patient $i \in \mathcal{N}$, $\beta_{ik}(\xi)$ is a binary parameter that denotes the caregiver with the patient's highest preference. In addition, $\beta_{ik}(\xi)$ can use for sharing information for patients with ongoing care in different weeks and as a result, it guarantees continuity of care between different weeks of serving patients. $o(\xi)$ is the maximum allowable number of inconsistency between patients' preferences and assigned caregivers.

The required number of temporary caregivers is determined before revealing the deterministic values of the uncertain parameters, and therefore it is considered as the first-stage variable in our proposed two-stage stochastic programming model. After determining the required number of caregivers, the scheduling, assignment, and routing decisions are determined under different scenarios. The required number of visits by each patient is scheduled during the planning horizon (one week), while $\mathcal{D}$ ($\mathcal{D} = \{1, 2, \ldots, d\}$) is the set of days in a week.

In this paper, the assignment of patients to a caregiver, scheduling, and routing are considered as the second-stage decision variables. In addition, it is assumed that caregivers start and end their routes from their homes. Therefore, selecting temporary caregivers to provide services will affect routing costs. After revealing the uncertain parameters, a group of patients with different required services is assigned to each selected caregiver for visiting. To ensure a balanced workload for each caregiver, the maximum number of patients assigned to caregivers is minimized in the objective function. The problem is illustrated in Fig. 2. In this figure, there are 3 available permanent and 2 temporary caregivers with different types of skills, and 15 patients are categorized into 3 classes. As shown in Fig. 2 one temporary caregiver is required to visit patients. The illustration in Fig. 2 is for a specific scenario in a working day. Therefore, as shown in this figure, visits of patients 1, 2, 4, 6, 7, 9, 12, 14, and 15 are scheduled on this particular day under the specific scenario and other patients are visited on other days of the week. The assumptions of the proposed model are stated as follows:

- HHC centers schedule a meeting to evaluate the required care and services by a patient. Based on the obtained information, patients are partitioned into different classes. Therefore, the classes of patients are unknown before this meeting and defined as an uncertain parameter. Also, they need an uncertain number of visits in their planning horizon, based on their classes. These uncertain parameters appear in a scenario-based scheme. For more explanation, the classes of patients, the required number of visits, and service time for two scenarios by considering 10 patients are reported in Table 2. In this example, we assume that patients are partitioned into three classes. As reported in the table, patients 5, 6, 7, 8, and 9 are in the first class under the first scenario and one visit should be scheduled for them during a week. However,

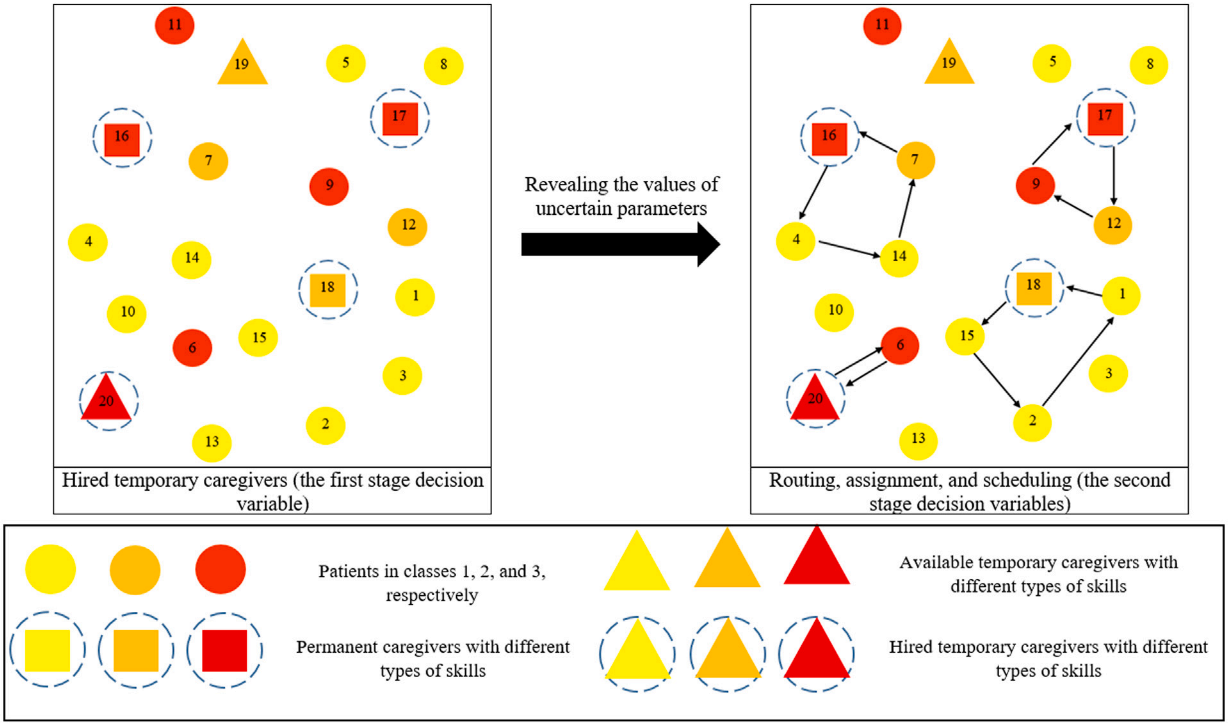


Fig. 2. A schematic view of the problem consisting of 5 caregivers and 15 patients. The color determines the level of skills and red and yellow are the highest and lowest level of skills, respectively.

Table 2
The classes of patients, required number of visits, and service time under scenarios.

Patient	Scenario 1			Scenario 2		
	The classes of patients	λ_i	st_i	The classes of patients	λ_i	st_i
1	2	2	67	2	2	88
2	3	3	101	1	1	20
3	2	2	31	1	1	21
4	3	3	77	3	3	89
5	1	1	24	1	1	31
6	1	1	22	3	3	88
7	1	1	36	2	2	70
8	1	1	41	2	2	31
9	1	1	18	1	1	18
10	2	2	63	2	2	67

in the second scenario, the classes of the mentioned patients are 1, 3, 2, 2, and 1, respectively. Therefore, one, three, two, two, and one visits should be scheduled during a week for those patients under the second scenario. In this example, it is assumed that the care of patients 1, 4, 5, 9, and 10 are continued from the previous weeks, therefore, their classes are known in all scenarios.

- Service time for each patient is unknown in advance and its deterministic value is revealed after visiting the patient. Therefore, service time for each patient is considered as a stochastic parameter and appears in a scenario-based scheme. Servicing times for 10 patients under two scenarios are reported in Table 2.
- Each caregiver can serve a limited number of patients. That means, the working time of each caregiver is less than or equal to a predetermined time.
- The time between two consecutive visits to each patient must be greater than a predefined value.
- The start and end points of each caregiver's route are his/her home.
- At most one visit can be scheduled for each patient per day.
- Under time is allowed for the caregivers in the proposed model but over time is not permitted.

The notations are defined consistently with previous papers, such as that by Nikzad et al. (2020), where possible. Other sets, parameters, and decision variables of the proposed model are as follows.

Sets:

$\mathcal{N}_c^k$	Set of the residential node of the kth caregiver (The start and end points of the route corresponding to the kth caregiver).
Ξ	Set of scenarios, $\Xi = \{1, 2, \dots, S\}$.

Parameters:

α	Balancing factor related to the maximum number of patients assigned to caregivers.
c_{ij}	Travel cost between the ith and the jth $(i, j \in \mathcal{V})$ nodes.
tt_{ij}	Travel time between the ith and the jth $(i, j \in \mathcal{V})$ nodes.
M_{max}	A large number, $M_{max} = N + K$.
T_{max}	A large number, $T_{max} = \max_i(b_i)$.
$p(\xi)$	Probability associated with scenario ξ $(\xi \in \Xi)$.

Decision variables:

$x_{ijkd}(\xi)$	1 if caregiver k $(k \in \mathcal{K})$ travels along (i, j) $(i, j \in \mathcal{V})$ in time period d $(d \in \mathcal{D})$ under scenario ξ $(\xi \in \Xi)$, 0 otherwise.
$u_{ikd}(\xi)$	1 if node i $(i \in \mathcal{V})$ is visited in the kth $(k \in \mathcal{K})$ route (corresponding to the kth caregiver) in time period d $(d \in \mathcal{D})$ under scenario $(\xi \in \Xi)$, 0 otherwise.
$z_{ik}(\xi)$	1 if patient i $(i \in \mathcal{N})$ is assigned to caregiver k $(k \in \mathcal{K})$ under scenario $(\xi \in \Xi)$, 0 otherwise.
y_k	1 if caregiver k $(k \in \mathcal{K})$ gets involved to serve the patients, 0 otherwise.
$\rho_i(\xi)$	Number of caregivers assigned to the ith $(i \in \mathcal{N})$ patient under scenario ξ $(\xi \in \Xi)$.

$$\min \sum_{k\in\mathcal{K}} h_k y_k + \sum_{\xi\in\Xi}\sum_{k\in\mathcal{K}}\sum_{d\in\mathcal{D}}\sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}} p(\xi)c_{ij}x_{ijkd}(\xi) + \sum_{i\in\mathcal{N}}\sum_{\xi\in\Xi} p(\xi)\theta_i\rho_i(\xi) + \alpha\sum_{\xi\in\Xi} p(\xi)\nu(\xi) \tag{1}$$

$$\text{s.t.: } y_k = 1, \quad \forall k \in \mathcal{K}_P \tag{2}$$

$$\sum_{i\in\mathcal{N}}\sum_{d\in\mathcal{D}} u_{ikd}(\xi) \geq y_k, \quad \forall k \in \mathcal{K}, \xi \in \Xi \tag{3}$$

$$\sum_{i\in\mathcal{N}} u_{ikd}(\xi) \leq M_{\max} u_{jkd}(\xi), \quad \forall k \in \mathcal{K}, d \in \mathcal{D}, j \in \mathcal{N}_c^k, \xi \in \Xi \tag{4}$$

$$\sum_{k\in\mathcal{K}}\sum_{d\in\mathcal{D}} u_{ikd}(\xi)\sum_{f\in\mathcal{F}}\sum_{a\in\mathcal{A}} r_{ifa}(\xi)q_{kfa} = \lambda_i(\xi), \quad \forall i \in \mathcal{N}, \xi \in \Xi \tag{5}$$

$$u_{ikd}(\xi) \leq y_k, \quad \forall i \in \mathcal{V}, k \in \mathcal{K}, d \in \mathcal{D}, \xi \in \Xi \tag{6}$$

$$\sum_{j\in\mathcal{V}} x_{ijkd}(\xi) = u_{ikd}(\xi), \quad \forall i \in \mathcal{N}, k \in \mathcal{K}, d \in \mathcal{D}, \xi \in \Xi \tag{7}$$

$$\sum_{j\in\mathcal{N}} x_{ijkd}(\xi) = u_{ikd}(\xi), \quad \forall i \in \mathcal{N}_c^k, k \in \mathcal{K}, d \in \mathcal{D}, \xi \in \Xi \tag{8}$$

$$\sum_{j\in\mathcal{N}} x_{jikd}(\xi) = u_{ikd}(\xi), \quad \forall i \in \mathcal{N}_c^k, k \in \mathcal{K}, d \in \mathcal{D}, \xi \in \Xi \tag{9}$$

$$\sum_{i\in\mathcal{V}, i\neq j} x_{ijkd}(\xi) = \sum_{i\in\mathcal{V} i\neq j} x_{jikd}(\xi), \quad \forall d \in \mathcal{D}, j \in \mathcal{N}, k \in \mathcal{K}, \xi \in \Xi \tag{10}$$

$$\sum_{j\in\mathcal{V} j\neq i}\sum_{k\in\mathcal{K}} x_{ijkd}(\xi) \leq 1, \quad \forall d \in \mathcal{D}, i \in \mathcal{V}, \xi \in \Xi \tag{11}$$

$$\sum_{i\in\mathcal{N}} st_i(\xi)u_{ikd}(\xi) \leq \varphi, \quad \forall d \in \mathcal{D}, k \in \mathcal{K}, \xi \in \Xi \tag{12}$$

$$t_{jkd}(\xi) \geq t_{ikd}(\xi) + st_i(\xi) + tt_{ij} - T_{max}\left(1 - x_{ijkd}(\xi)\right), \quad \forall d \in \mathcal{D}, i \in \mathcal{N}, j \in \mathcal{V}, k \in \mathcal{K}, \xi \in \Xi \tag{13}$$

$$t_{ikd}(\xi) \leq b_i, \quad \forall d \in \mathcal{D}, i \in \mathcal{N}, k \in \mathcal{K}, \xi \in \Xi \tag{14}$$

$$z_{ik}(\xi) \geq u_{ikd}(\xi), \quad \forall i \in \mathcal{N}, k \in \mathcal{K}, d \in \mathcal{D}, \xi \in \Xi \tag{15}$$

$$\rho_i(\xi) \geq \sum_{k\in\mathcal{K}} z_{ik}(\xi), \quad \forall i \in \mathcal{N}, \xi \in \Xi \tag{16}$$

$$\nu(\xi) \geq \sum_{i\in\mathcal{N}} u_{ikd}(\xi), \quad \forall k \in \mathcal{K}, d \in \mathcal{D}, \xi \in \Xi \tag{17}$$

$$\sum_{d'\in\{d+1,\dots,d+\delta_{fa}\}}\sum_{k\in\mathcal{K}} u_{ikd'}(\xi)r_{ifa}(\xi) \leq 1 - \sum_{k\in\mathcal{K}} u_{ikd}(\xi)r_{ifa}(\xi),$$
$$\forall i \in \mathcal{N}, d \in \{1, 2, \dots, D - \delta_{fa}\}, f \in \mathcal{F}, a \in \mathcal{A}, \xi \in \Xi \tag{18}$$

$$\sum_{d\in\{D-\delta_{fa}+1,\dots,D\}}\sum_{k\in\mathcal{K}} u_{ikd}(\xi)r_{ifa}(\xi) \leq 1, \quad \forall i \in \mathcal{N}, f \in \mathcal{F}, a \in \mathcal{A}, \xi \in \Xi \tag{19}$$

Box I.

$t_{ikd}(\xi)$ The time that servicing starts at the ith ($i \in \mathcal{N}$) patient by the kth ($k \in \mathcal{K}$) caregiver in time period d ($d \in \mathcal{D}$) under scenario ξ ($\xi \in \Xi$).

$\nu(\xi)$ Maximum number of patients visited by caregivers in different time periods under scenario ξ ($\xi \in \Xi$).

The proposed two-stage stochastic model for staff-dimensioning problem is presented here (see the equation in Box I):

$$\sum_{k\in\mathcal{K}}\sum_{i\in\mathcal{N}} z_{ik}(\xi)(1-\beta_{ik}(\xi)) \leq o(\xi), \quad \forall \xi \in \Xi \tag{20}$$

$$y_k,\ u_{ikd}(\xi),\ x_{ijkd}(\xi) \in \{0,1\}, \quad \forall d \in \mathcal{D}, i \in \mathcal{V}, j \in \mathcal{V}, k \in \mathcal{K}, \xi \in \Xi \tag{21}$$

$$z_{ik}(\xi), \in \{0,1\}, \quad \forall i \in \mathcal{N}, k \in \mathcal{K}, \xi \in \Xi \tag{22}$$

$$t_{ikd}(\xi) \geq 0, \quad \forall d \in \mathcal{D}, i \in \mathcal{V}, k \in \mathcal{K}, \xi \in \Xi \tag{23}$$

$$\rho_i(\xi), \nu(\xi) \in \mathbb{Z}^{0+}, \quad \forall i \in \mathcal{N}, \xi \in \Xi \tag{24}$$

The objective function minimizes the total hiring cost of caregivers, the expected value of transportation and number of caregivers assigned to patients, and the maximum number of patients assigned to a caregiver in a day. The number of caregivers assigned to each patient is minimized in the third part of the objective function, meaning that this term tries to assign the same caregiver to a patient and considers continuity of care. Constraint (2) states that all permanent caregivers are available. Constraint (3) guarantees that if a caregiver is used, at least one patient must be assigned to him/her. Constraint (4) states that the kth caregiver can visit patients in day d under scenario ξ if this caregiver leaves his/her home to serve patients. Constraint (5) ensures that compatible caregivers are assigned to each patient and the required visits for each patient are scheduled during the time period. Constraint (6) states that the ith patient can be assigned only to one of the employed caregivers. Constraint (7) assures that each patient should be served by the corresponding route of the assigned caregiver in the day d. Constraints (8) and (9) ensure that the kth route (corresponding to the kth caregiver) is started and ended at the location of the kth caregiver. Constraint (10) is a flow conservation constraint. Constraint (11) guarantees that each patient is visited by one caregiver at most in each day. Constraint (12) guarantees that the working time

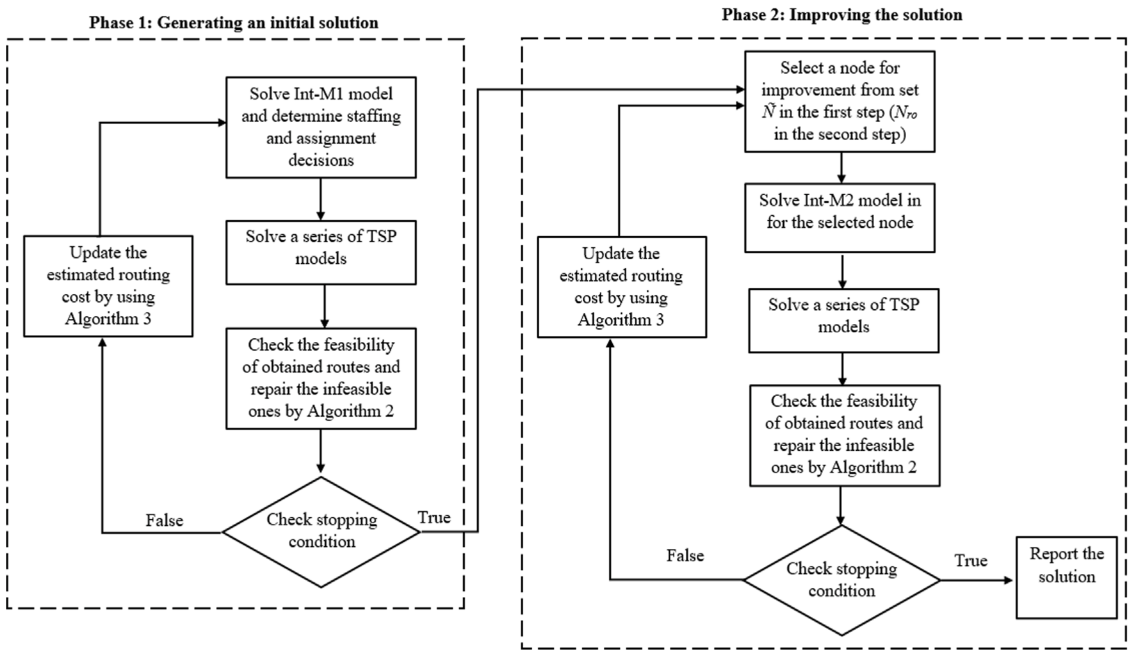


Fig. 3. An overview of the proposed matheuristic algorithm containing two phases for generating initial solutions and their improvement.

of each caregiver is less than the maximum permitted working time. Constraints (13) and (14) are sub-tour elimination and time-window constraints, respectively. Constraints (15) and (16) determine the number of caregivers assigned to each patient. Constraint (17) calculates the maximum number of patients assigned to a caregiver in per day under each scenario. Constraints (18) and (19) guarantee a minimum required duration between two consecutive visits to each patient. Constraint (20) guarantees that the number of assigned caregivers without considering patients' highest preference is less than a predefined value. Constraints (21)–(24) determine domains of the decision variables.

4. Two-phase matheuristic algorithm

In this paper, a two-phase matheuristic algorithm is developed for the staff-dimensioning problem. Archetti and Speranza (2014) named this type of matheuristic algorithm as improvement heuristic and showed that MILP model can be used for generating an initial solution or improving the obtained solution in this type of matheuristic algorithm. In this paper, the problem is decomposed into several sub-problems according to the type of decisions to generate an initial solution. An initial solution is generated with a combination of an intermediate model (Int-M1) and a series of TSPs solved for each caregiver per day under each scenario (TSP(k,d,ξ)). In this phase, these models are solved iteratively to find a good quality initial solution. In each iteration, the required number of temporary caregivers and the assignment of patients to permanent and temporary caregivers per day under each scenario are determined by solving the Int-M1 model according to an estimation of routing cost. Then a series of TSPs is solved for each caregiver per day under each scenario to obtain routing decisions according to the assignment decisions achieved from the Int-M1 model in the current iteration. The estimated routing cost is updated based on the obtained routes to use in the Int-M1 model again in the next iteration. After convergence of the first phase, in the second phase, the obtained solutions for assignment and routing variables are improved by using a fix-and-optimize procedure. In each iteration of this phase, a patient is selected and her/his corresponding routing cost is improved while assignment and routing decisions for other patients are fixed. The second phase includes two steps, in the first step, routing decisions are improved for all patients. In this step, patients are selected for improvement based on a predefined order. In the second step, the routing decisions are improved for some patients with the highest potential of improvement. Pseudocode for the matheuristic algorithm is presented in Algorithm 1 also an overview of the proposed matheuristic algorithm is shown in Fig. 3.

The additional sets and parameters in the matheuristic algorithm are as follows.

Sets:

$\mathcal{N}^*_{kd}(\xi)$	Set of patients visited by the kth ($k \in \mathcal{K}$) caregiver in time period d ($d \in \mathcal{D}$) under scenario ξ ($\xi \in \Xi$).
$\tilde{\mathcal{N}}$	Set of potential nodes for improvement in the first step of improvement phase in the proposed matheuristic algorithm.
$\mathcal{N}_{ro}(\xi)$	Set of potential nodes for improvement under scenario ξ ($\xi \in \Xi$) in the second step of improvement phase in the proposed matheuristic algorithm.

Parameters:

$\Gamma^m_{ikd}(\xi)$	Estimated routing cost if the ith ($i \in \mathcal{N}$) patient is visited by the kth ($k \in \mathcal{K}$) caregiver in day d ($d \in \mathcal{D}$) under scenario ξ ($\xi \in \Xi$) in the mth iteration.
$route_{kd}(\xi)$	The obtained tour for the kth ($k \in \mathcal{K}$) caregiver at day d ($d \in \mathcal{D}$) under scenario ξ ($\xi \in \Xi$).
cc_{ij}	Cost of inserting node i in the jth position for the current route. It is used in Algorithm 3 to calculate the estimated routing cost.
w_{ij}	Cost of changing the position of the ith patient from the current position to the jth position in its corresponding route. It is used in Algorithm 2 for making the solution to be feasible.
F^m	Objective value of Int-M1 model in the mth iteration.
$Maxiter$	Maximum number of iterations that the first phase of the proposed matheuristic algorithm is done.
ε	Maximum allowable difference between the obtained objective values of Int-M1 model in two consecutive iterations.

Input: Parameters of the model, $Maxiter$, ε
Phase 1: Generating an initial solution
condition ← false, $m \leftarrow 1$, $F^0 \leftarrow \infty$;
Use Eq. (25) to calculate $\Gamma^m_{ikd}(\xi)$;
while *condition is false* **do**
  Optimize the Int-M1 model and determine the values of decision variables ($\hat{y}^m_k$, $\hat{u}^m_{ikd}(\xi)$, $\hat{z}^m_{ik}(\xi)$, $\hat{\rho}^m_i(\xi)$, $\hat{v}^m(\xi)$) and objective value (F^m). Also obtain the sets $\mathcal{N}^*_{kd}(\xi)$;
  Determine the routing decisions by solving all the TSP(k, d, ξ) based on the results obtained from optimizing the Int-M1 model;
  Check the feasibility of obtained routes and repair the infeasible routes by Algorithm 2;
  if $|F^m - F^{m-1}| \leq \varepsilon$ *or* $m = Maxiter$ **then**
    condition ← true ;
  else
    $m = m + 1$ and use Algorithm 3 to update the $\Gamma^m_{ikd}(\xi)$;
  end
end
$\tilde{y}_k \leftarrow \hat{y}^m_k$, $\tilde{u}_{ikd}(\xi) \leftarrow \hat{u}^m_{ikd}(\xi)$, $\tilde{\rho}_i(\xi) \leftarrow \hat{\rho}^m_i(\xi)$, $\tilde{v}(\xi) \leftarrow \hat{v}^m(\xi)$, $\tilde{z}_{ik}(\xi) \leftarrow \hat{z}^m_{ik}(\xi)$, ;
Phase 2: Improving the solution
Step 1 of improvement phase
for $\xi \in \Xi$ **do**
  $\tilde{N} \leftarrow \mathcal{N}$, $\Gamma^1_{ikd}(\xi) = \Gamma^m_{ikd}(\xi)$, $m \leftarrow 1$;
  while $|\tilde{N}| \neq \emptyset$ **do**
    Select a node ($\tilde{i}$) from the set $\tilde{\mathcal{N}}$ and update the set $\tilde{\mathcal{N}}$ by removing $\tilde{i}$ from this set;
    Optimize the Int-M2 model and determine the values of decision variables ($\hat{u}^m_{\tilde{i}kd}(\xi)$, $\hat{z}^m_{\tilde{i}k}(\xi)$, $\hat{\rho}^m_{\tilde{i}}(\xi)$), $\hat{v}^m(\xi)$) and update the sets $\mathcal{N}^*_{kd}(\xi)$;
    $\tilde{u}_{ikd}(\xi) \leftarrow \hat{u}^m_{ikd}(\xi)$, $\tilde{z}_{ik}(\xi) \leftarrow \hat{z}^m_{ik}(\xi)$, $\tilde{\rho}_i(\xi) \leftarrow \hat{\rho}^m_i(\xi)$, $\tilde{v}(\xi) \leftarrow \hat{v}^m(\xi)$;
    Optimize the TSP(k, d, ξ) model to determine the optimal routing decision based on the solution obtained from Int-M2 and determine $route_{kd}(\xi)$;
    Check the feasibility of obtained routes and repair the infeasible routes by Algorithm 2;
    $m = m + 1$;
    Use Algorithm 3 to update the value of parameter $\Gamma^m_{ikd}(\xi)$;
  end
end
Step 2 of improvement phase
for $\xi \in \Xi$ **do**
  Define the set $N_{ro}(\xi)$ which its elements are potentially improvable, $\Gamma^1_{ikd}(\xi) \leftarrow \Gamma^m_{ikd}(\xi)$, $m \leftarrow 1$;
  while $|N_{ro}(\xi)| \neq \emptyset$ **do**
    Select a node ($\tilde{i}$) from the set $N_{ro}(\xi)$ with largest required number of visits and in case of existing alternatives, select among them with smallest value of $\min_{k,d}\{\Gamma^m_{ikd}(\xi)\}$;
    Remove $\tilde{i}$ from the set $N_{ro}(\xi)$;
    Optimize the Int-M2 model and determine the values of decision variables ($\hat{u}^m_{ikd}(\xi)$, $\hat{z}^m_{ik}(\xi)$, $\hat{\rho}^m_i(\xi)$, $\hat{v}^m(\xi)$) and update the sets $\mathcal{N}^*_{kd}(\xi)$;
    $\tilde{u}_{ikd}(\xi) \leftarrow \hat{u}^m_{ikd}(\xi)$, $\tilde{z}_{ik}(\xi) \leftarrow \hat{z}^m_{ik}(\xi)$, $\tilde{\rho}_i(\xi) \leftarrow \hat{\rho}^m_i(\xi)$, $\tilde{v}(\xi) \leftarrow \hat{v}^m(\xi)$;
    Optimize the TSP(k, d, ξ) model to determine the optimal routing decision based on the solution obtained from the Int-M2 model and determine $route_{kd}(\xi)$;
    Check the feasibility of obtained routes and repair the infeasible routes by Algorithm 2;
    $m = m + 1$;
    Use Algorithm 3 to update the value of parameter $\Gamma^m_{ikd}(\xi)$;
  end
end
The objective value (F_{best}) is obtained according to $\tilde{y}_k$, $\tilde{\rho}_i(\xi)$, $\tilde{v}(\xi)$, and $route_{kd}(\xi)$;
Result: F_{best}, $route_{kd}(\xi)$

Algorithm 1: The framework of the proposed matheuristic algorithm.

Decision variables:

l_{ij} 1 if the ith patient is assigned to the jth position in the route, 0 otherwise.

4.1. Phase 1: Generating an initial solution

The aim of this phase is to determine good quality initial solutions for the decision variables. In this phase, an iterative procedure is proposed by considering an estimation of transportation costs. A brief overview of this phase is as follows:

- Calculate the estimated routing cost.
- Solve the Int-M1 model to obtain decisions related to staff dimensioning, scheduling, and assigning patients to caregivers according to the estimated routing cost.
- Solve a series of TSPs for each caregiver per day under each scenario (TSP(k, d, ξ)) to determine routing decisions according to the assignment decisions obtained from Int-M1 model.
- Repair the infeasible routes achieved from TSP(k, d, ξ) by a proposed procedure (refer to Algorithm 2).

As mentioned above, the estimated routing cost ($\Gamma^m_{ikd}(\xi)$) should be calculated first in this phase. The estimated cost in the first iteration of this phase is calculated by using the following equation.

$$\Gamma^1_{ikd}(\xi) = 2c_{i(N+k)} \qquad \forall i \in \mathcal{N},\ k \in \mathcal{K},\ d \in \mathcal{D},\ \xi \in \Xi \tag{25}$$

In other iterations, the estimated cost is updated based on the obtained routes from the previous iteration. Then the required number of caregivers, and the assignment of patients to caregivers are determined by solving the Int-M1 model. This model is used to generate a good quality initial solutions for all decisions except routing. In each iteration, the required number of temporary caregivers with different types of skills and the assignment of patients to caregivers are determined using this model. The proposed Int-M1 model is presented as follows.

$$(\textbf{Int}-\textbf{M1}) \ \min \sum_{k\in\mathcal{K}} h_k y_k + \sum_{\xi\in\Xi}\sum_{k\in\mathcal{K}}\sum_{d\in\mathcal{D}}\sum_{i\in\mathcal{N}} p(\xi)\Gamma^m_{ikd}(\xi)u_{ikd}(\xi) + \theta_i \sum_{i\in\mathcal{N}}\sum_{\xi\in\Xi} p(\xi)\rho_i(\xi) + \alpha\sum_{\xi\in\Xi} p(\xi)v(\xi) \tag{26}$$

s.t.: (2), (3), (5), (6), (12), (15), (16)–(20), *and*

$$\sum_{k\in\mathcal{K}} u_{ikd}(\xi) \leq 1, \qquad \forall i \in \mathcal{N},\ d \in \mathcal{D},\ \xi \in \Xi \tag{27}$$

$$u_{ikd}(\xi) \in \{0,1\}, \qquad \forall d \in \mathcal{D},\ i,j \in \mathcal{V},\ k \in \mathcal{K},\ \xi \in \Xi \tag{28}$$

$$z_{ik}(\xi) \in \{0,1\}, \qquad \forall i \in \mathcal{N},\ k \in \mathcal{K},\ \xi \in \Xi \tag{29}$$

$$\rho_i(\xi), v(\xi) \in \mathbb{Z}^{0+}, \qquad \forall i \in \mathcal{N},\ \xi \in \Xi \tag{30}$$

The objective function minimizes the hiring cost of caregivers, the expected value of transportation, the number of caregivers assigned to patients, and the maximum number of patients assigned to a caregiver in a time-period. Constraint (27) guarantees that each patient is visited by one caregiver at most in each time period. The values of the decision variables y_k, $u_{ikd}(\xi)$, $\rho_{ik}(\xi)$ and $v(\xi)$ (which are indicated by $\hat{y}_k$, $\hat{u}_{ikd}(\xi)$, $\hat{\rho}_{ik}(\xi)$ and $\hat{v}(\xi)$, respectively) are achieved by optimizing this model. Also, the set $\mathcal{N}^*_{kd}(\xi)$ is defined as a set of patients assigned to the kth caregiver at day d under each scenario ξ in the current iteration. This set is updated in each iteration based on the value of the decision variable $u_{ikd}(\xi)$, which is obtained by solving the int-M1 model.

To determine routing decisions according to the values obtained from the Int-M1 model, a series of TSPs are solved. Thus, the route related to the kth caregiver in day d under different scenarios in iteration m is obtained by solving a series of TSPs (k, d, ξ) for each caregiver in each day under each scenario, based on the values of the decision variables $\hat{y}^m_k$ and $\hat{u}^m_{ikd}(\xi)$ obtained from the Int-M1 model. $route_{kd}(\xi)$ is the route obtained from TSPs (k, d, ξ). The time-window constraints are not considered in the TSPs, therefore the tours obtained from them can be infeasible. So, the feasibility of the routes obtained from the TSPs is examined and repaired.

For repairing an infeasible obtained tour, a procedure is proposed in this paper and its pseudocode is provided in Algorithm 2. In this procedure, visiting sequence of patients in infeasible routes is changed to obtain feasible ones. At first inserting cost related to adding of each patient in different positions of infeasible route is calculated. For this reason, the ith patient is selected and removed from $route_{kd}(\xi)$, while the visiting sequence of other patients in the route is not changed. Then the costs corresponding to inserting the ith patient in different positions on $route_{kd}(\xi)$ ($w_{ij} \quad \forall j \in \{2,\dots,|route_{kd}(\xi)|-1\}$) are calculated, where $|route_{kd}(\xi)|$ shows the number of nodes on the mentioned route. If the route obtained from visiting the ith patient in the jth position of $route_{kd}(\xi)$ is feasible, the corresponding cost (w_{ij}) is calculated by $w_{ij} = c_{i_{j^-}i} + c_{ii_{j^+}} - c_{i_{j^-}i_{j^+}}$, where i_{j^-} and i_{j^+} denote the predecessor and successor nodes of patient i when it is inserted in the jth position of $route_{kd}(\xi)$, respectively. However, if the achieved route is infeasible, w_{ij} is set to be a large number. Then Int-MF model is solved to obtain a feasible visiting sequence with minimum cost for each infeasible route. The Int-MF model is as follows.

$$(\mathbf{Int-MF(k,d,\xi)}) \min \sum_{i\in\mathcal{N}^*_{kd}(\xi)} \sum_{j\in\{2,\dots,|route_{kd}(\xi)|-1\}} w_{ij} l_{ij} \tag{31}$$

$$\text{s.t.:} \sum_{j\in\{2,\dots,|route_{kd}(\xi)|-1\}} l_{ij} = 1, \ \forall i \in \mathcal{N}^*_{kd}(\xi) \tag{32}$$

$$\sum_{i\in\mathcal{N}^*_{kd}(\xi)} l_{ij} = 1, \qquad \forall j \in \{2,\dots,|route_{kd}(\xi)|-1\} \tag{33}$$

$$l_{ij} \in \{0,1\}, \qquad \forall i \in \mathcal{N}^*_{kd}(\xi), j \in \{2,\dots,|route_{kd}(\xi)|-1\} \tag{34}$$

The objective function (31) minimizes the cost of insertion. Constraint (32) guarantees that each patient in $route_{kd}(\xi)$ is assigned to one position. Constraint (33) ensures that each position can be allocated to exactly one patient. Constraint (34) determines the domain of variables. $route_{kd}(\xi)$ is updated based on the solution obtained from Int-MF and this procedure is continued until feasible routes are determined. Then, if the stopping condition of Algorithm 1 is met, the solution obtained in the last iteration is considered as the initial solution for the whole algorithm. However, if the stopping condition is not met, then the estimated routing cost must be updated based on the solution obtained for routing decisions using the next proposed procedure.

The procedure proposed for updating estimated routing cost is illustrated in Algorithm 3. In this algorithm, the estimated costs corresponding to caregivers who are not selected to serve patients ($\hat{y}_k = 0$) are calculated by Eq. (25). Also, for the selected caregivers ($\hat{y}_k = 1$), the estimated cost for days and scenarios when caregivers do not visit any patient is similarly obtained from Eq. (25). In the case that the ith patient is served by the kth caregiver on day d under scenario ξ, the estimated cost is calculated by $c_{i^-i} + c_{ii^+} - c_{i^-i^+}$, where, i^- and i^+ denote the predecessor and successor nodes of patient i, respectively.

When a patient is not visited on $route_{kd}(\xi)$, the corresponding estimated cost is determined based on the least cost of inserting this patient into this route. To calculate the least insertion cost, the patient is added to different positions on a current route, and the corresponding costs (cc_{ij}) are determined. If time-window constraints are not violated for all nodes on $route_{kd}(\xi)$ after adding the ith patient to the jth position of the route, the insertion cost corresponding to the jth position is calculated by $cc_{ij} = c_{i_{j^-}i} + c_{ii_{j^+}} - c_{i_{j^-}i_{j^+}}$, but if time-window constraints are violated in the jth position, cc_{ij} is increased to a large number ($BigM$). Thus the solutions obtained for assignment and scheduling in all iterations except the first iteration are always feasible.

For further clarification, an example is given to explain the proposed procedure for calculating $\Gamma^m_{ikd}(\xi)$ when node i is not visited by the kth caregiver on day d under scenario ξ. In this example, 5 caregivers with different types of skills are available and 4 caregivers are selected to serve 15 patients. The first caregiver is not selected to provide

```
Input: route_kd(ξ), N*_kd(ξ), b_i, st_i, tt_ij
for k ∈ K do
    if ŷ_k^{m-1} = 1 then
        for d ∈ D do
            for ξ ∈ Ξ do
                if route_kd(ξ) is infeasible then
                    condition3 ← false;
                    while condition3 if false do
                        for i ∈ N*_kd(ξ) do
                            Remove the i^th patient from route_kd(ξ)
                             when visiting sequence of other nodes
                             are constant;
                            for j ∈ {2, ..., |route_kd(ξ)| − 1} do
                                Insert the i^th node in the j^th position
                                 in the route_kd(ξ);
                                if the obtained route is feasible then
                                    w_ij = c_{i_j- i} + c_{i i_j+} − c_{i_j- i_j+};
                                else
                                    w_ij = BigM;
                                end
                            end
                        end
                        Solve Int-MF model and obtain route_kd(ξ);
                        if route_kd(ξ) is feasible then
                            condition3 ← true;
                        end
                    end
                end
            end
        end
    end
end
Result: route_kd(ξ)
```

Algorithm 2: The framework for making the obtained route from TSPs to be feasible.

caring services for patients ($\hat{y}_1 = 0$). The obtained tours from the TSPs based on the solutions calculated for the intermediate model (Int-M1 or Int-M2) for selected caregivers in the second day under scenario 1 are reported in Table 3. Based on the solutions obtained from intermediate models, under the first scenario patients 1, 2, and 15 are assigned to the second caregiver and patients 12 and 9 are visited by the fourth caregiver on the second day of the week. Their corresponding routes are 17-15-2-1-17 and 19-12-9-19, respectively. Under the first scenario, the estimated routing cost corresponding to patient 12 if he/she is visited by the second caregiver on the second day of the week in the mth iteration ($\Gamma^m_{12,2,2}(1)$) is determined based on insertion costs corresponding to different positions on the route of the second caregiver ($cc_{ij}, \quad \forall j \in \{1,2,\dots,|route_{2,2}(1)|\}$).

As reported in Table 3, $route_{2,2}(1)$ includes three patients, so four positions can be considered to insert a new node (i) to this route. On the other hand, the number of potential positions for inserting a new patient into the route corresponding to the kth caregiver on day d under scenario ξ is equal to $|route_{kd}(\xi)|-1$. After inserting a new patient into the jth potential position and generating a new route, arrival times are calculated for all nodes in the new route ($i' \in route_{kd}(\xi) \cup \{i\}$). Fig. 4 shows the potential locations and corresponding insertion costs for inserting patient 12 into the route of the second caregiver, based on the routes reported in Table 3.

4.2. Phase 2: Improving the obtained solutions

In this phase, a fix-and-optimize procedure is used to improve assignment and routing decisions. The improvement is achieved in two steps and the decision to select caregivers is determined based on the solution obtained from the first phase. In each iteration of this phase, a patient is selected and the corresponding assignment and routing decisions are improved when the staff dimensioning and routing decisions for other patients are fixed. In the first step, the routing decisions

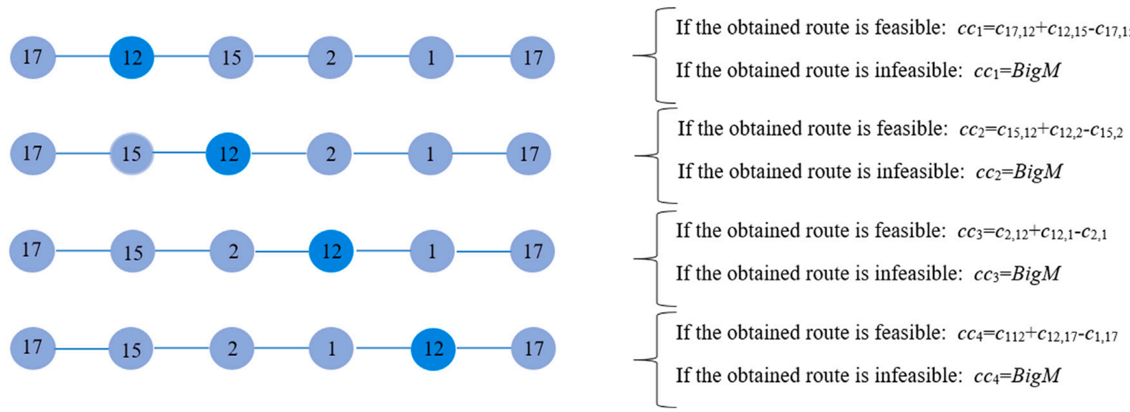


Fig. 4. Calculating the costs corresponding to insert node 12 into different positions (cc_{ij} $\forall j = 1:4$) of the route $17-15-2-1-17$. For example, $cc_{12,1}$ shows the cost related to add node 12 in the first position (between nodes 17 and 15). These positions are defined as the locations between each sequential nodes.

Input: $route_{kd}(\xi)$, $N^*_{kd}(\xi)$, c_{ij}, b_i, st_i, tt_{ij}
for $k \in \mathcal{K}$ **do**
  if $\hat{y}^{m-1}_k = 1$ **then**
    for $d \in \mathcal{D}$ **do**
      for $\xi \in \Xi$ **do**
        if *$route_{kd}(\xi)$ is not empty* **then**
          for $i \in \mathcal{N}$ **do**
            if $i \in N^*_{kd}$ **then**
              $\Gamma^m_{ikd}(\xi) = c_{i^-i} + c_{ii^+} - c_{i^-i^+}$;
            else
              for $j = 1 : |N^*_{kd}(\xi)| + 1$ **do**
                Calculate the cost corresponding to insert the i^{th} node into the j^{th} (cc_{ij}) position on $route_{kd}(\xi)$;
                Calculate the arrival time at node i and other nodes in this route when it is added in the j^{th} position on $route_{kd}(\xi)$;
                if *Arrival time for at least one node is greater than its corresponding time-window* **then**
                  $cc_{ij} = BigM$;
                else
                  $cc_{ij} = c_{i_j^- i} + c_{ii_j^+} - c_{i_j^- i_j^+}$
                end
              end
              $\Gamma^m_{ikd}(\xi) = \min_j \{cc_{ij}\}$;
            end
          end
        else
          $\Gamma^m_{ikd}(\xi) = 2c_{i(N+k)}$
        end
      end
    end
  else
    for $d \in \mathcal{D}$ **do**
      for $\xi \in \Xi$ **do**
        $\Gamma^m_{ikd}(\xi) = 2c_{i(N+k)}$
      end
    end
  end
end
Result: $\Gamma^m_{ikd}(\xi)$

Algorithm 3: The framework for updating parameter $\Gamma^m_{ikd}(\xi)$ in iteration m.

corresponding to all patients are improved. While, in the second step, routing decisions are improved for some patients with a higher possible improvement. The improvement phase of the proposed algorithm has been adopted from Nikzad et al. (2020), however, there are some differences that are listed here. The improvement phase of this study contains two steps and the existence of two steps will increase the chance of more improvement while it was only one step in the above mentioned study. Moreover, in the first step of the improvement phase of the algorithm, a method has been proposed to order the patients to be examined in a fix-and-optimize procedure while there is no such ordering scheme in the above-mentioned study. The proposed Int-M2 model has been adopted from the Int-M4 model of above mentioned study with required differences related to the nature of the problem. A brief overview of the steps in the second phase is described as follows:

Table 3
Allocated routes for selected caregivers in the illustrative example.

Selected caregivers	Corresponding routes
2	17-15-2-1-17
3	18-6-18
4	19-12-9-19
5	20-7-14-4-20

- Select a patient for improvement.
- Calculate cost corresponding to insert the selected patient to each caregiver's route per day under each scenario ($\Gamma^m_{ikd}(\xi)$).
- Solve the Int-M2 model for the selected patient to obtain decisions related to scheduling and assigning the selected patient to caregivers according to the estimated routing cost.
- Solve a series of TSPs for each caregiver per day under each scenario (TSP(k, d, ξ)) to determine routing decisions.
- Repair the infeasible routes achieved from TSP(k, d, ξ) by developing a procedure (refer Algorithm 2).

In each iteration of the first step, a patient ($\tilde{i}$) is selected for improvement. Let $\tilde{\mathcal{N}}$ be the set of potential patients for improvement and include all patients in the first iteration of this step. Selection of the patient for improvement from set $\tilde{\mathcal{N}}$ is based on two factors. First, the patient with the largest required number of visits is selected and, in case of existing alternatives, a patient with the smallest value of $\min_{k,d}\{\Gamma^m_{ikd}(\xi)\}$ is selected. Then the set $\tilde{\mathcal{N}}$ is updated by removing $\tilde{i}$ from it. The assignment and scheduling decisions for patient $\tilde{i}$ are optimized by solving the Int-M2 model when these decisions are fixed for other patients ($j \in \mathcal{V}\backslash\{\tilde{i}\}$). Improving the scheduling and assignment decisions and as a result, decreasing the routing cost is the aim of the Int-M2 model. The Int-M2 model is as follows.

$$(\textbf{Int}-\textbf{M2}(\xi,\tilde{\textbf{i}})) \qquad \min \sum_{k\in\boldsymbol{\mathcal{K}}}\sum_{d\in\boldsymbol{D}} \Gamma^m_{\tilde{i}kd}(\xi)u_{\tilde{i}kd}(\xi) + \theta\rho_{\tilde{i}}(\xi) + \alpha\nu(\xi) \tag{35}$$

s.t.: (3), (5), (6), (12), (15), (16)–(20), (27) *and*

$$y_k = \tilde{y}_k, \qquad \forall k \in \boldsymbol{\mathcal{K}} \tag{36}$$

$$u_{ikd}(\xi) = \tilde{u}_{ikd}(\xi), \qquad \forall i \in \mathcal{V}\backslash\{\tilde{i}\},\ k \in \mathcal{K},\ d \in \mathcal{D} \tag{37}$$

$$z_{ik}(\xi) = \tilde{z}_{ik}(\xi), \qquad \forall i \in \mathcal{N}\backslash\{\tilde{i}\},\ k \in \mathcal{K},\ d \in \mathcal{D} \tag{38}$$

$$u_{ikd}(\xi), z_{ik}(\xi)\ y_k \in \{0,1\}, \qquad \forall d \in \mathcal{D},\ i,j \in \mathcal{V},\ k \in \mathcal{K} \tag{39}$$

$$\rho_i(\xi), v(\xi) \in \mathbb{Z}^{0+}, \qquad \forall i \in \mathcal{N} \tag{40}$$

Constraints (37) and (38) ensure that the assignment decisions are fixed for all nodes except the one selected in the Int-M2 model. The values of variables $u_{ikd}(\xi)$, $z_{ik}(\xi)$, and $\rho_i(\xi)$ for the selected patient are determined in each iteration. Int-M2 is solved for all patients under each scenario, therefore, it is optimized $N \times S$ times in the first step. In this phase, a series of TSPs are solved for each caregiver on each day under each scenario to achieve routing decisions based on the obtained assignment solutions from Int-M2. Then the feasibility of obtained routes is checked and infeasible routes are repaired. If the set of potential patients for improvement is not empty, $\Gamma^m_{ikd}(\xi)$ is updated based on the routing decisions from Algorithm 3.

Further improvement is explored in the second step for the patients whose corresponding assignment decisions were altered in the first step. To obtain the set of mentioned patients ($\mathcal{N}_{ro}(\xi)$), first, $k^*_i(\xi)$ and $d^*_i(\xi)$ corresponding to minimum values of $\Gamma^m_{ikd}(\xi)$ are determined for each patient under each scenario by $(k^*_i(\xi), d^*_i(\xi)) = \arg\min_{k,d}\{\Gamma^m_{ikd}(\xi)\}$. Then the set $\mathcal{N}_{ro}(\xi)$ is defined to include the ith patient if this patient is not visited by caregiver $k^*_i(\xi)$ on day $d^*_i(\xi)$ under scenario ξ. On the other hand, the decision variable $u_{ik^*_i(\xi),d^*_i(\xi)}$ is equal to zero for these patients. Then, as in the previous step, a patient is selected from set $\mathcal{N}_{ro}(\xi)$, and the Int-M2 model is optimized to improve routing and assignment decisions. Based on the solution obtained from Int-M2, TSP models are solved to achieve routing decisions. Then infeasible routes are repaired and is updated based on the routing decisions from Algorithm 3. In the first step, routing and assignment decisions related to all patients are improved, whereas in the second step improvement is explored for some patients as to the possibility of reducing costs by changing their caregiver assignment. Therefore, the sets of potential patients for improvement are different in these steps. This step stops when the set $\mathcal{N}_{ro}(\xi)$ is empty.

5. Computation results

In this section, the efficiency and validity of the algorithm are investigated through numerical experiments. Instances are generated based on 5 instances for urban areas by considering 75 nodes reported by the Austrian Red Cross (Fikar and Hirsch, 2015). The upper bounds of time windows and the intervals to define the classes of patients are also taken from these instances. Some parameters used in the proposed model have no values in the dataset, so the values for parameters are shown in Table 4. The generated instances can be found in https://github.com/bashirimahdi/StaffDimensionning-for-HHC. The computations are performed on a 3.5 GHz Workstation with 32 GB RAM and 6 cores operating on Windows 10 (64-bit) using Julia software. The CPLEX 12.7.1 solver is used for solving all intermediate models and TSPs in the proposed matheuristic algorithm. Also, a limited computational time (1000 s) is considered for the Int-M1 model in the first phase of the matheuristic algorithm.

5.1. Scenario selection scheme

Increasing the number of scenarios in the two-stage stochastic programming models means that considerably more computational time is required to solve them. Nevertheless, when the number of possible scenarios is large, a sample of scenarios can be selected as representative instead of using all scenarios. Scenario selection has a significant impact on the accuracy of obtained solutions. Here we develop a scenario generation procedure for our problem. In this procedure, the objective value obtained from the proposed matheuristic algorithm under each scenario is used to generate a sample of scenarios, which is then partitioned into S clusters with the k-means algorithm (MacQueen, 1967) based on their obtained objective values, where S is the size of samples. For example, scenarios are partitioned into 10 clusters when a sample with 10 scenarios is required. Then, a scenario is selected randomly from each cluster.

Table 4
Additional parameters' values in numerical instances.

Parameters	Corresponding values or distribution
N	10, 20, 30, 40, 50, 60, 70
K	5 10, 15
S	10, 20
D	7
F	1
A	3
xx_i	Discrete uniform [1, 100]
yy_i	Discrete uniform [1, 100]
c_{ij}	$\sqrt{(xx_i - xx_j)^2 + (yy_i - yy_j)^2}$
tt_{ij}	c_{ij}
$\lambda_i(\xi)$	Discrete uniform [1, 3]
e	uniform[0.5, 1.5]
$st_i(\xi)$	$\lambda_i(\xi) \times 30 \times e$
φ	300, 480
θ_i	100
α	40

To investigate the efficiency of the proposed scenario generation procedure, we use the in-sample index proposed by Kaut and Wallace (2003). This index indicates how similar the solutions obtained from different samples of scenarios are. To examine the in-sample stability of the proposed approach, 10 different samples are selected and the coefficients of variation (CV) for the objective function values of these samples are reported in Fig. 5. The stability of the proposed procedure is examined for the different numbers of patients by considering 10 and 20 scenarios in each sample. As shown in this figure, considering 20 scenarios generated by the proposed scenario generation results in a small CV value, showing the stability of the model even using a relatively small number of scenarios (e.g. 20). As shown in the figure, this ratio is less than 0.1 when 20 scenarios are considered for different numbers of patients.

5.2. Sensitivity analysis of the proposed model

In this section, the efficiency of the proposed mathematical model is examined and the effect of different parameters on the obtained values of decision variables is investigated. First, due to the importance of workload balancing for caregiver satisfaction, the effects of parameter α on the routing cost and average value of the maximum number of patients assigned to caregivers are analyzed. The results obtained are shown in Fig. 6. As seen there, the maximum number of patients assigned to caregivers decreases with larger values of α, however, as expected the corresponding routing cost increases, confirming the validity of the proposed model.

With a fixed number of caregivers, increasing α reduces the maximum number of patients allocated to caregivers at the cost of increasing travel distance (i.e. travel cost). Fig. 6 shows that although travel costs will increase, the maximum number of patients allocated to caregivers will decrease, which makes it more reasonable. This analysis confirms the applicability of the model in different situations for healthcare services, as well as its validity. This finding is linked to that of Porzio et al. (2020), who considered a home healthcare system serving cancer patients during the COVID-19 pandemic. They scheduled twice-weekly visits for patients with moderate symptoms, while patients with severe symptoms were visited every day of the week. The classes of patients

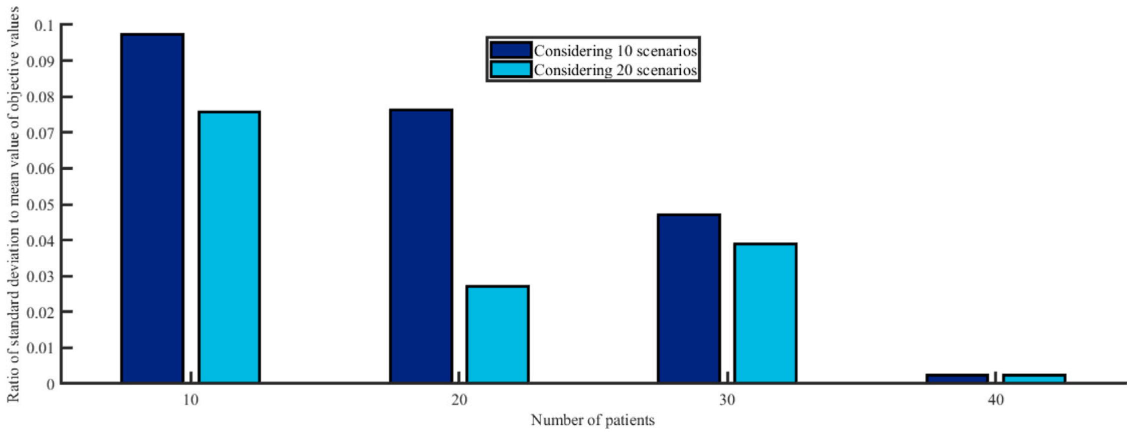
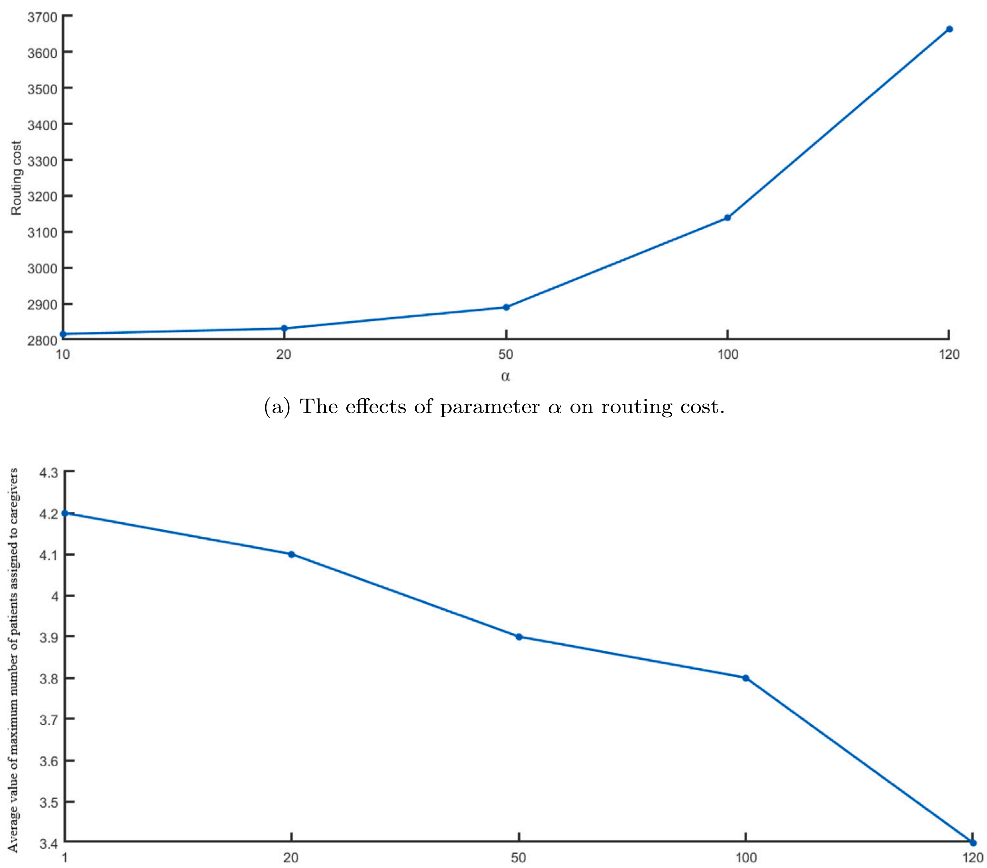


Fig. 5. Analyzing the stability of solutions obtained from the different numbers of scenarios.



(a) The effects of parameter α on routing cost.

(b) The effects of parameter α on maximum number of patients assigned to caregivers

Fig. 6. Analyzing the effect of importance parameter related to the maximum number of allocated patients by considering 30 patients, 10 caregivers, and 10 scenarios.

and the required number of visits in this analysis are determined based on data reported by Porzio et al. (2020), so it can be concluded that the proposed model can also be used efficiently under pandemic conditions to consider the importance of decreasing the number of contacts.

5.3. Examining the efficiency of the proposed matheuristic algorithm

The efficiency of the proposed matheuristic algorithm to solve the staff dimensioning problem in HHC services of different sizes is investigated in this section. The results obtained from the proposed matheuristic algorithm are compared with the results given by the CPLEX as well as an approach based on a combination of the Progressive Hedging and the Frank–Wolfe algorithm (FW-PH) (Boland et al., 2018) for small-size instances (when $N \times K \times S$ is less than or equal to 4000) in Table 5. The optimal value of objective function (Obj_C) and the best found objective function values achieved by matheuristic algorithm (Obj_{Math}) are reported in this table. Also, the lower bound provided by the CPLEX in 10000 s (LB_C) is reported in this table. Obj_{FW-PH} reported the objective value obtained by fixing the first stage decision variables obtained by the FW-PH in the proposed model. The gap between Obj_{Math} and Obj_C is calculated by $Gap_{Math-C} = (Obj_{Math} - Obj_C)/Obj_C \times 100\%$ and is reported in column

Table 5
Computation results for different numbers of patients, caregivers, and scenarios in five instances. Index Math and C denote the obtained results by our proposed matheuristic and the CPLEX, respectively.

N	K	K_1	S	ins	Obj_C	LB_C	$Time_C$ (s)	Obj_{Math}	$Time_{Math}$ (s)	Obj_{FW-PH}	$Time_{FW-PH}$ (s)	Gap_{Math-C}(%)
10	5	2	10	1	8522.92	8314.09	10000	8576.30	46.60	8513.44	3414.18	0.63
				2	7163.98	7034.23	10000	7230.64	42.16	7163.78	587.56	0.93
				3	7058.23	6910.34	10000	7177.53	42.83	7051.05	1018.79	1.69
				4	7115.75	6921.66	10000	7158.60	43.11	7068.48	2993.13	0.60
				5	8582.24	8334.16	10000	8622.34	56.91	8532.76	3058.72	0.47
			20	1	8697.37	8311.25	10000	8662.14	68.36	8555.15	8652.75	**−0.41**
				2	8646.70	8312.71	10000	8642.80	80.24	8531.73	11580.56	**−0.05**
				3	8696.62	8317.65	10000	8671.87	65.22	8544.33	26880.498	**−0.28**
				4	8594.61	8295.35	10000	8600.09	70.55	8510.46	13653.94	0.06
				5	8624.81	8291.76	10000	8582.60	68.58	8498.70	4541.34	**−0.49**
20	10	5	10	1	17758.77	12223.53	10000	13665.67	80.83			**−23.05**
				2	14766.37	13045.12	10000	13725.73	88.00			**−7.05**
				3	19249.46	12034.43	10000	12547.89	72.64			**−34.81**
				4	18876.33	13072.92	10000	14960.40	149.82			**−20.75**
				5	16189.58	11983.00	10000	12519.55	76.75			**−22.67**
			20	1	16836.47	11711.16	10000	13699.73	160.83			**−18.63**
				2	16256.38	11875.67	10000	12543.25	176.26			**−22.84**
				3	17871.24	12870.89	10000	13691.36	151.20			**−23.39**
				4	18034.27	12859.03	10000	13693.36	173.27			**−24.07**
				5	19309.86	11741.68	10000	13654.87	203.18			**−29.29**
30	10	5	10	1	24706.35	13856.84	10000	16597.75	788.60			**−32.82**
				2	23080.92	15126.22	10000	16536.02	265.18			**−28.36**
				3	22713.83	13689.62	10000	15467.80	468.07			**−31.90**
				4	22979.55	15898.15	10000	17731.28	355.18			**−22.84**
				5	22851.27	15940.49	10000	17635.56	331.97			**−22.82**
Average									165.05			−14.49

13 of Table 5. Computational times for CPLEX and the matheuristic algorithm are also shown in the table.

The results indicate that CPLEX cannot find optimal solutions in any cases where computational time is limited to 10000 s. Nevertheless, the proposed matheuristic algorithm can obtain better solutions than CPLEX in most cases when the average computational time of the matheuristic algorithm is 165.05 s. Also, the average gap between Obj_{Math} and Obj_C is −14.49%. It is concluded that the proposed matheuristic algorithm can achieve better quality solutions in a reasonable time for small-size instances. The FW-PH algorithm obtained better solutions than CPLEX and matheuristic algorithm by considering 10 patients. However, the FW-PH algorithm is not an efficient solution method for this problem comparing to the proposed matheuristic. The reason is that the computational time of the FW-PH is large even for small-size instances as the proposed model cannot be solved in a reasonable time for each scenario. especially for instances with more than 10 patients the computational time of the FW-PH is more than 86400 s (24 h) while the proposed matheuristic algorithm can solve instances with 85 nodes in a reasonable computational time.

The efficiency of the matheuristic algorithm for medium-size and large-size instances is illustrated in Tables 6 and 7, respectively. The effect of the number of scenarios on the efficiency of the proposed matheuristic algorithm is analyzed for medium-size instances (when $N \times K \times S$ is less than 12000). The results are presented in Table 6. The results show that the CPLEX cannot find any feasible solutions for the model in 10000 s, while the matheuristic algorithm can find good quality solutions in 1322.57 s on average. Also, the lower bounds obtained from the CPLEX solver in 10000 s are reported in this table. As reported in this table, the CPLEX cannot find proper lower bounds for the instances with 70 patients. Therefore, it can be concluded that the developed matheuristic algorithm performs properly for medium-size instances, too.

Results of analyzing the effect of increasing the number of scenarios on the efficiency of the matheuristic algorithm for large-size instances (when $N \times K \times S$ is less than or equal to 21000) are reported in Table 7. The CPLEX cannot find feasible solutions for the model in 10000 s for any instance, while the proposed algorithm finds proper solutions for all cases in 3814.95 s on average. Also, the CPLEX cannot find a lower bound for the instances with more than 40 patients because of memory limitation. The results demonstrate the algorithm's efficiency to find good quality solutions for large-size instances.

Table 6
Computation results for different numbers of caregivers and scenarios for medium instances. Index Math and C denote the obtained results by our proposed matheuristic and the full model.

N	K	K_1	S	ins	Obj_C	LB_C	Obj_{Math}	$Time_{Math}$ (s)
30	10	5	20	1	–	13045.79	17592.17	1222.06
				2	–	15468.92	17730.64	1562.86
				3	–	12973.22	16429.62	1731.04
				4	–	13105.49	17673.95	1654.67
				5	–	12964.39	16511.33	1031.34
40	15	8	10	1	–	17741.70	19205.62	267.62
				2	–	17781.33	19218.41	457.61
				3	–	17642.83	19145.33	280.01
				4	–	17729.57	19144.26	283.82
				5	–	17772.57	19165.41	268.03
50	15	8	10	1	–	18874.72	20807.04	705.7
				2	–	18819.35	20855.32	814.74
				3	–	13200.00	20847.37	789.34
				4	–	18796.86	20764.73	575.49
				5	–	13200.00	21126.08	1227.86
60	15	8	10	1	–	19911.21	22326.43	1545.49
				2	–	13200.00	23224.87	1456.24
				3	–	19861.79	22356.46	2060.71
				4	–	19837.36	22711.49	992.47
				5	–	19911.99	22485.98	1579.19
70	15	8	10	1	–	13200.00	23206.94	2045.89
				2	–	13200.00	23488.71	1810.65
				3	–	13200.00	24360.28	3018.12
				4	–	13200.00	23305.32	2912.33
				5	–	13200.00	23354.59	2770.93
Average								1322.57

The effect of increasing the number of scenarios on the computational times of the matheuristic algorithm is examined and the results are shown in Fig. 7. The mean values (mean) and standard deviations

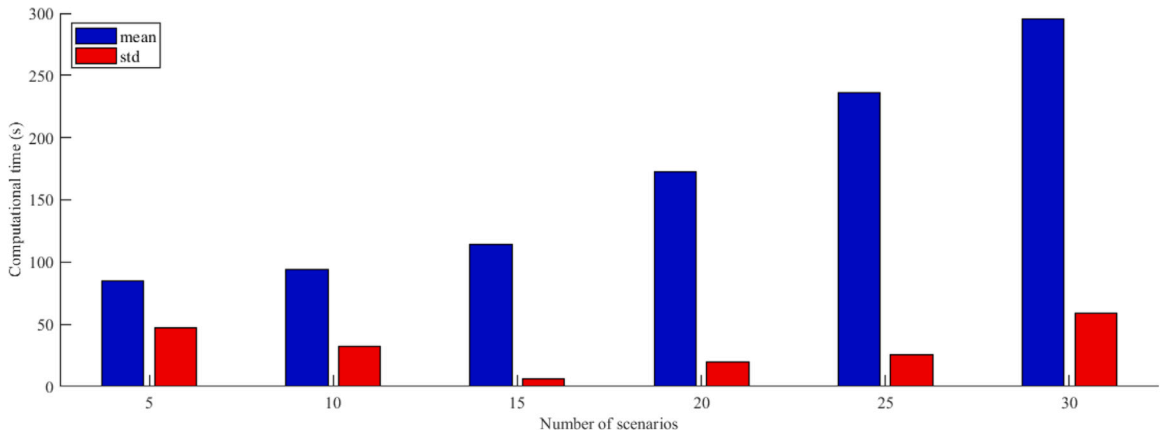


Fig. 7. Analyses for the effect of increasing the number of scenarios on the computational time of the proposed matheuristic algorithm.

(std) of the matheuristic algorithm's computational times with 20 patients, 5 permanent, and 5 temporary caregivers, as well as 7 days under the different number of scenarios are shown in this figure. As mentioned previously, increasing the number of scenarios will increase the computational time of the two-stage stochastic model. In the proposed matheuristic algorithm, the two-stage stochastic model is used in the first phase of the algorithm and in subsequent phases the problems are decomposed based on scenarios by fixing the first-stage decision variables. Also, in the first phase of the proposed algorithm, the number of variables dependent on the scenarios is reduced by removing the routing decisions. Therefore, the computational time of the algorithm is not increased greatly by increasing the number of scenarios.

5.4. Evaluating the efficiency of the proposed matheuristic algorithm under different benchmarks

The proposed model in this study focused on staff dimensioning, assignment, scheduling, and routing decisions simultaneously. The location routing problem with time windows is the most similar classic problem to the problem studied in this paper. Therefore, we considered the developed benchmarks for the location routing problem with time windows by Ponboon et al. (2016) to examine the efficiency of the proposed matheuristic algorithm. The modified benchmarks are designed based on Ponboon et al. (2016)[3] are used in this study. Because of existing differences between the location routing problem and the problem of this study, some changes have been done in the proposed matheuristic algorithm

- Some changes were done in Int-M1 and Int-M2 models. For example, in the proposed staff dimensioning problem, it is assumed that one vehicle (caregiver) is available in each depot (residential location of caregiver) but in these benchmarks some vehicles are available in depots which leads to the Int-M1 and Int-M2 models modification. Also, the constraints about the capacity of depots and vehicles were considered.
- The proposed procedure to select nodes for improvement in the first step of the improvement phase in the developed matheuristic algorithm cannot be used for the location routing problem. Because this procedure is designed based on the required number of visits by patients but in the location routing problem each node is visited once. Therefore, a new procedure is designed based on the amount of saving cost only. So, the node with greater value of saving cost is selected for improvement in each iteration of this step.

Table 7
Computation results for different numbers of caregivers and scenarios for large instances. Index Math and C denote the obtained results by our proposed matheuristic and the full model.

N	K	K_1	S	ins	Obj_C	LB_C	Obj_{Math}	$Time_{Math}$ (s)
40	15	8	20	1	–	17780.97	19252.93	1896.12
				2	–	17689.39	19147.08	2109.08
				3	–	17783.91	19181.08	1711.57
				4	–	17778.08	19184.63	1611.48
				5	–	17831.54	19254.90	1970.35
50	15	8	20	1	–	*	20838.15	2851.21
				2	–	*	21243.83	6250.36
				3	–	*	22432.40	6043.98
				4	–	*	23064.79	6140.13
				5	–	*	20886.54	6182.67
60	15	8	20	1	–	*	22278.02	4702.92
				2	–	*	22370.02	6540.66
				3	–	*	23202.65	3612.57
				4	–	*	23521.46	2547.39
				5	–	*	22271.06	4235.64
70	15	8	20	1	–	*	23210.30	2855.49
				2	–	*	24225.89	4791.01
				3	–	*	23251.42	4469.98
				4	–	*	23528.69	3513.34
				5	–	*	23228.04	2263.03
Average								3814.95

*: not capable of compiling.

The obtained results from applying the proposed algorithm for the LRP benchmarks are reported in Table 8. The reported results show that the proposed algorithm has a satisfactory performance for LRP benchmarks.

6. Conclusion

Staff dimensioning is a critical problem in home healthcare systems and significantly impacts their efficiency. In this paper, staff dimensioning, routing, and assignment decisions are considered simultaneously. Staff dimensioning is usually a long-term decision. Nevertheless, routing and assignment decisions in HHC are determined each day. Therefore, two types of caregivers are considered in this study. The decision about the required number of permanent caregivers is a predefined input parameter, however the decision about the required number of

[3] http://prodhonc.free.fr/Instances/instances_us.htm

Table 8
Examining the efficiency of the proposed matheuristic algorithm for LRP benchmarks.

Instance LRP optimal	Solution	Obj_{math}	Gap (%)
coordGaspelle	424.90	429.73	1.14
coordGaspelle2	585.11	586.70	0.27
coordGaspelle3	512.10	519.08	1.30
coordGaspelle4	562.2	562.2	0.00
coordGaspelle5	504.33	518.16	2.74
coordGaspelle6	460.37	481.66	4.62
coordMin27	3062.02	3065.24	0.10
Average			1.45

temporary caregivers is considered as a tactical decision. In this study, the number of required temporary caregivers is minimized over a one week period.

Staff dimensioning is done in the planning phase, where there are several uncertainties, such as the classes of patients, the required number of visits, and the service times. In addition, considering uncertainty in the planning horizon makes decision-makers more capable of dealing with uncertainty in a cost-effective manner. In this study, an integrated two-stage stochastic model was proposed to formulate the staff dimensioning problem.

This study also considered compatibility between patients' required types of care and caregivers' skills, continuity of care, patient preferences on caregiver assignment, caregiver workload balancing, and visiting regulations. These factors were taken into account to make more realistic decisions and to increase patients' and caregivers' satisfaction. Although considering uncertainty increases the complexity of the proposed model, a matheuristic algorithm was developed to find high quality solutions in a reasonable time. This algorithm consists of two phases, where a good solution is generated in the first phase by using an intermediate model and a series of TSP models. The obtained solution was then improved by using a fix-and-optimize procedure based on optimizing another intermediate model and a series of TSP models.

The efficiency of the proposed algorithm was investigated using extensive numerical studies. The results confirmed efficiency of the proposed matheuristic algorithm for small, medium, and large-size instances. For small-size instances, this algorithm found solutions with an average gap of −14.49% from the best solutions obtained by the CPLEX (in 10000 s), while the optimal solutions were not found in any cases of 25 instances by the CPLEX. Also, the CPLEX cannot find any feasible solutions for medium and large-size instances in 10000 s, while the matheuristic algorithm can find proper solutions in reasonable times of 1322.57 and 3814.95 s for medium and large-size instances, respectively.

The proposed model offers the opportunity for decision-makers to compromise between travel costs and the number of patients to be visited by the same caregivers and consider the continuity of care. This is vital in pandemic situations such as COVID-19 pandemic when it is preferred that the same caregiver to be allocated to a set of patients to reduce the risk of spreading the virus as much as possible, even though it may increase total cost because caregivers might need to travel longer distances. Investigation of all relevant factors in home healthcare staff dimensioning with a focus on a home care service case study might be a valuable direction for the future study.

CRediT authorship contribution statement

Erfaneh Nikzad: Data curation, Writing – original draft, Software, Validation, Verification, Investigation, Formal analysis. **Mahdi Bashiri:** Conceptualization, Methodology, Validation, Writing – review & editing, Formal analysis, Supervision, Investigation, Project administration. **Babak Abbasi:** Visualization, Investigation, Formal analysis, Writing – review & editing.

Acknowledgments

Authors would like to thank anonymous reviewers for their provided valuable comments which lead to a significant improvement of the paper.

References

Allaoua, H., Borne, S., Létocart, L., Calvo, R.W., 2013. A matheuristic approach for solving a home health care problem. Electron. Notes Discrete Math. 41, 471–478.
Andersen, A.R., Nielsen, B.F., Reinhardt, L.B., Stidsen, T.R., 2019. Staff optimization for time-dependent acute patient flow. European J. Oper. Res. 272 (1), 94–105.
Archetti, C., Speranza, M.G., 2014. A survey on matheuristics for routing problems. EURO J. Comput. Optim. 2 (4), 223–246.
Bahadori-Chinibelagh, S., Fathollahi-Fard, A.M., Hajiaghaei-Keshteli, M., 2019. Two constructive algorithms to address a multi-depot home healthcare routing problem. IETE J. Res. 1–7.
Bertazzi, L., Coelho, L.C., De Maio, A., Laganà, D., 2019. A matheuristic algorithm for the multi-depot inventory routing problem. Transp. Res. E Logist. Transp. Rev. 122, 524–544.
Birge, J.R., Louveaux, F., 2011. Introduction to Stochastic Programming. Springer Science & Business Media.
Boland, N., Christiansen, J., Dandurand, B., Eberhard, A., Linderoth, J., Luedtke, J., Oliveira, F., 2018. Combining progressive hedging with a Frank–Wolfe method to compute Lagrangian dual bounds in stochastic mixed-integer programming. SIAM J. Optim. 28 (2), 1312–1336.
Braekers, K., Hartl, R.F., Parragh, S.N., Tricoire, F., 2016. A bi-objective home care scheduling problem: Analyzing the trade-off between costs and client inconvenience. European J. Oper. Res. 248 (2), 428–443.
Campelo, P., Neves-Moreira, F., Amorim, P., Almada-Lobo, B., 2019. Consistent vehicle routing problem with service level agreements: A case study in the pharmaceutical distribution sector. European J. Oper. Res. 273 (1), 131–145.
Cappanera, P., Scutellà, M.G., Nervi, F., Galli, L., 2018. Demand uncertainty in robust home care optimization. Omega 80, 95–110.
Carello, G., Lanzarone, E., 2014. A cardinality-constrained robust model for the assignment problem in home care services. European J. Oper. Res. 236 (2), 748–762.
Cinar, A., Salman, F.S., Bozkaya, B., 2021. Prioritized single nurse routing and scheduling for home healthcare services. European J. Oper. Res. 289 (3), 867–878.
Cissé, M., Yalçındağ, S., Kergosien, Y., Şahin, E., Lenté, C., Matta, A., 2017. OR problems related to Home Health Care: A review of relevant routing and scheduling problems. Oper. Res. Health Care 13, 1–22.
Decerle, J., Grunder, O., El Hassani, A.H., Barakat, O., 2018. A memetic algorithm for a home health care routing and scheduling problem. Oper. Res. Health Care 16, 59–71.
Dorneles, Á.P., de Araújo, O.C., Buriol, L.S., 2014. A fix-and-optimize heuristic for the high school timetabling problem. Comput. Oper. Res. 52, 29–38.
Fathollahi-Fard, A.M., Ahmadi, A., Goodarzian, F., Cheikhrouhou, N., 2020a. A bi-objective home healthcare routing and scheduling problem considering patients' satisfaction in a fuzzy environment. Appl. Soft Comput. 93, 106385.
Fathollahi-Fard, A.M., Ahmadi, A., Karimi, B., 2020b. A robust optimization for a home healthcare routing and scheduling problem considering greenhouse gas emissions and stochastic travel and service times. In: Green Transportation and New Advances in Vehicle Routing Problems. Springer, pp. 43–73.
Fikar, C., Hirsch, P., 2015. A matheuristic for routing real-world home service transport systems facilitating walking. J. Clean. Prod. 105, 300–310.
Fikar, C., Hirsch, P., 2017. Home health care routing and scheduling: A review. Comput. Oper. Res. 77, 86–95.
Grenouilleau, F., Legrain, A., Lahrichi, N., Rousseau, L.-M., 2019. A set partitioning heuristic for the home health care routing and scheduling problem. European J. Oper. Res. 275 (1), 295–303.
Hashemi Doulabi, H., Pesant, G., Rousseau, L.-M., 2020. Vehicle routing problems with synchronized visits and stochastic travel and service times: Applications in healthcare. Transp. Sci. 54 (4), 1053–1072.
Hernandez, F., Gendreau, M., Jabali, O., Rei, W., 2019. A local branching matheuristic for the multi-vehicle routing problem with stochastic demands. J. Heuristics 25 (2), 215–245.
Kaut, M., Wallace, S.W., 2003. Evaluation of Scenario-Generation Methods for Stochastic Programming. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät ....
Lanzarone, E., Matta, A., 2014. Robust nurse-to-patient assignment in home care services to minimize overtimes under continuity of care. Oper. Res. Health Care 3 (2), 48–58.
Li, Y., Chu, F., Chu, C., Zhu, Z., 2019. An efficient three-level heuristic for the large-scaled multi-product production routing problem with outsourcing. European J. Oper. Res. 272 (3), 914–927.
Lin, C.-C., Hung, L.-P., Liu, W.-Y., Tsai, M.-C., 2018. Jointly rostering, routing, and rerostering for home health care services: A harmony search approach with genetic, saturation, inheritance, and immigrant schemes. Comput. Ind. Eng. 115, 151–166.

Lindahl, M., Sørensen, M., Stidsen, T.R., 2018. A fix-and-optimize matheuristic for university timetabling. J. Heuristics 24 (4), 645–665.

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Vol. 1. No. 14. Oakland, CA, USA, pp. 281–297.

Miranda, P.L., Cordeau, J.-F., Ferreira, D., Jans, R., Morabito, R., 2018. A decomposition heuristic for a rich production routing problem. Comput. Oper. Res. 98, 211–230.

Moussavi, S.E., Mahdjoub, M., Grunder, O., 2019. A matheuristic approach to the integration of worker assignment and vehicle routing problems: Application to home healthcare scheduling. Expert Syst. Appl. 125, 317–332.

Neves-Moreira, F., Almada-Lobo, B., Cordeau, J.-F., Guimarães, L., Jans, R., 2019. Solving a large multi-product production-routing problem with delivery time windows. Omega 86, 154–172.

Nguyen, T.V.L., Toklu, N.E., Montemanni, R., 2015. Matheuristic optimization for robust home health care services. In: Proc. ICAOR. p. 2.

Nikzad, E., Bashiri, M., Abbasi, B., 2020. A matheuristic algorithm for stochastic home health care planning. European J. Oper. Res. 288 (3), 753–774.

Ponboon, S., Qureshi, A.G., Taniguchi, E., 2016. Branch-and-price algorithm for the location-routing problem with time windows. Transp. Res. E Logist. Transp. Rev. 86, 1–19.

Porzio, G., Cortellini, A., Bruera, E., Verna, L., Ravoni, G., Peris, F., Spinelli, G., 2020. Home care for cancer patients during COVID-19 pandemic: The double triage protocol. J. Pain Symptom Manage. 60 (1), e5–e7. http://dx.doi.org/10.1016/j.jpainsymman.2020.03.021.

Rodriguez, C., Garaix, T., Xie, X., Augusto, V., 2015. Staff dimensioning in homecare services with uncertain demands. Int. J. Prod. Res. 53 (24), 7396–7410.

Shi, Y., Boudouh, T., Grunder, O., Wang, D., 2018. Modeling and solving simultaneous delivery and pick-up problem with stochastic travel and service times in home health care. Expert Syst. Appl. 102, 218–233.

Solyalı, O., Süral, H., 2017. A multi-phase heuristic for the production routing problem. Comput. Oper. Res. 87, 114–124.

Vieira, B., Demirtas, D., van de Kamer, J.B., Hans, E.W., van Harten, W., 2018. A mathematical programming model for optimizing the staff allocation in radiotherapy under uncertain demand. European J. Oper. Res. 270 (2), 709–722.

Yalçındağ, S., Cappanera, P., Scutellà, M.G., Şahin, E., Matta, A., 2016. Pattern-based decompositions for human resource planning in home health care services. Comput. Oper. Res. 73, 12–26.

Yu, Q., Fang, K., Zhu, N., Ma, S., 2019. A matheuristic approach to the orienteering problem with service time dependent profits. European J. Oper. Res. 273 (2), 488–503.

Yuan, B., Liu, R., Jiang, Z., 2015. A branch-and-price algorithm for the home health care scheduling and routing problem with stochastic service times and skill requirements. Int. J. Prod. Res. 53 (24), 7450–7464.

Yuan, B., Liu, R., Jiang, Z., 2018. Daily scheduling of caregivers with stochastic times. Int. J. Prod. Res. 56 (9), 3245–3261.

Zhan, Y., Wang, Z., Wan, G., 2020. Home service routing and appointment scheduling with stochastic service times. European J. Oper. Res. http://dx.doi.org/10.1016/j.ejor.2019.07.009.