Computers & Operations Research 135 (2021) 105425

Contents lists available at ScienceDirect

Computers and Operations Research

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

An optimization-based approach for the healthcare districting under uncertainty

Sobhan Mostafayi Darmian [a], Mohammad Fattahi [b,*], Esmaeil Keyvanshokooh [c]

[a] Department of Industrial Engineering, University of Kurdistan, Sanandaj, Iran
[b] Newcastle Business School, Northumbria University, Newcastle Upon Tyne, United Kingdom
[c] Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, MI, USA

ARTICLE INFO

Keywords:
Healthcare services
Districting problem
Robust optimization
Mixed-integer programming
Graph theory
Graph-based genetic algorithm

ABSTRACT

In this paper, we address a districting problem motivated by a real-world case study to partition residential areas of a region for a healthcare system operation under interval demand uncertainty of health services. A mixed-integer programming model is presented based on the graph theory that enforces contiguity constraints on districts in addition to other corresponding practical criteria. Furthermore, robust optimization approaches are extended to address the existing uncertainty. To deal with the problem's computational intractability, an improved genetic algorithm is developed in accordance with the graph-based nature of the problem obtaining near-optimal solutions on large-sized instances. Extensive computational results are presented on a real-world case study and several randomly generated instances to evaluate the applicability of the models, performance of the presented robustness measure, and effectiveness of the solution approach. Furthermore, a hierarchical districting approach that enables decision makers to obtain districting decisions in various levels of health services is examined. Sensitivity analyses on main parameters are performed to derive some managerial insights that can help practitioners in providing suitable and homogeneous health services in a geographical area.

1. Introduction

There are several strategic and tactical planning contexts in which decision makers have to divide a geographical area (e.g., a city, state, country, etc.) into clusters or districts comprised of basic functional units (e.g., residential areas, customers, streets, zip codes, etc.). *Districting problems* aim at grouping basic units into districts under the jurisdiction of a decision maker in an optimized way, subject to several functional and topological constraints (Kalcsics and Ríos-Mercado, 2019). Districting applications include service management, such as healthcare management, emergency response planning, and municipal solid waste collection (see e.g., Benzarti et al., 2013; Lin and Kao, 2009; Ríos-Mercado and Bard, 2019; Mayorga et al., 2013); political organization (see e.g., Ricca et al., 2013; Ricca and Simeone, 2008), distribution planning (see e.g., Zhong et al., 2007; Lei et al., 2012), and commercial territory planning (see e.g., Salazar-Aguilar et al., 2011).

This research is motivated by our collaborations with Iran's Ministry of Health and Medical Education (MOHME) that desires to partition Iran's geographical area into a finite number of districts by using systematic approaches based on Operations Research techniques. Generally, a healthcare system operates in an extreme uncertain and variable environment, and health services should be consistently accessible and easily available over a geographical area with different population densities and characteristics (Harper et al., 2005). In such a condition, partitioning a geographical area, where a healthcare system functions, into a finite number of districts, can improve the accessibility to the healthcare system in all residential areas, health network management, medical logistics planning, and assignment of equipment to different areas. The healthcare services can be categorized into three main levels as *level I* that includes specialty and subspecialty medical/paramedical services, *level II* that should be provided by doctors and nurses in clinics and hospitals, and *level III* that consists of health centers such as emergency centers, treatment centers, health homes, and so on. In this study, we deal with the healthcare districting problem, and our goal is to develop an optimization-based framework for healthcare districting that is applicable for local, regional, and national healthcare planners. It is worth noting that depending on service levels, there are various criteria and concerns in their corresponding healthcare districting problem.

In this study, we focus on the healthcare districting problem at the

* Corresponding author at: Newcastle Business School, Northumbria University, Newcastle Upon Tyne, United Kingdom.
E-mail address: mohammad.fattahi@northumbria.ac.uk (M. Fattahi).

https://doi.org/10.1016/j.cor.2021.105425
Received 14 August 2019; Received in revised form 9 June 2021; Accepted 9 June 2021
Available online 15 June 2021
0305-0548/

first level, and our goal is to develop an optimization-based approach for this challenging problem. Although the healthcare districting is scarcely addressed in the literature, proper districting decisions: (1) enhance and facilitate the healthcare system management in the level of districts instead of a large geographical area; (2) maximize the balance among districts based on their capacities and demands; (3) improve the accessibility of patients to the healthcare system by reducing the total travel number of patients to other districts for accessing to the healthcare services. According to the literature (Harper et al., 2005; Steiner et al., 2015; Datta et al., 2013) and the requirements of our real-world case study, the main functional and topological constraints corresponding to healthcare districting problem at the first level of healthcare services include:

1. *The contiguity of districts*: The residential areas within each district should be adjacent to each other, i.e. a district is said to be contiguous if it is possible to travel between each two different basic units of the district without crossing the district's boundary (Kalcsics and Ríos-Mercado, 2019).
2. *The balance among districts*: Balance defines the desire for equal size districts in accordance with some context-dependent performance measures for districts (Kalcsics and Ríos-Mercado, 2019). In our problem, the districts' ability in serving the patients, based on their capacity and size of the population that requires the corresponding health services, is considered to define the balance criterion.
3. *Exclusive assignment*: Each residential area should belong to one and only one district.
4. *The maximum traveling time/distance through districts*: To have an efficient health network management through a district, which includes the reallocation of health workers, equipment, and health service capacities between healthcare providers, and the patients' accessibility to the health services, the maximum traveling time/distance between residential areas of a district should be less than a predetermined limit.

The above-mentioned requirements are considered for the healthcare districting at the first level of services in this paper. Based on World Health Organization, there exists a hierarchical structure among three levels of services and to address the districting at various levels, a framework for the hierarchical districting is examined. In other words, the districts obtained at each level should be partitioned in the next level of districting. It should be noted that the healthcare districting at the second and third levels might have special constraints and objectives those are not our concerns in this study and we have just proposed a hierarchical districting framework for this purpose.

Generally, there exist two critical challenges in developing a mixed-integer programming (MIP) model for the healthcare districting problem. First, the balance may be interpreted differently in various contexts, and there is not any specific criterion to define it. Further, incorporating a balance criterion into mathematical programs may result in a nonlinear and non-convex term and several studies consider lower and upper bounds for districts' size instead of obtaining perfectly balanced districts (Kalcsics and Ríos-Mercado, 2019). Secondly, presenting an MIP formulation to address the contiguity of districts is difficult and most of the presented mathematical models are not tractable in large-sized instances. Graph theory-based mathematical programs are proposed regarding the second challenge (see e.g., Duque et al., 2011), although they are not usually solvable by commercial solvers in large-sized instances. As another popular approach, a number of inequalities based on given neighborhood relations embedded into the model, which increases exponentially with the number of basic units. A few studies developed cutting plane-based methods based on their problem's structure (see e.g., Ríos-Mercado and López-Pérez, 2013) to handle the intractability in such a modeling approach.

This study proposes an MIP formulation, called the spanning tree-based model, and according to the graph-based nature of the problem, we enforce contiguity constraints on districts in addition to other above-mentioned criteria. The formulation uses the spanning tree idea that its essence is used by Ahuja et al. (1990), Yamada (2009), and Kim (2018) with a different modeling approach. Then, an enumeration-based heuristic approach is developed to solve the spanning tree-based model for all spanning trees.

Many different solution approaches have been proposed for the districting problem in the related literature. Due to the NP-hardness of the problem (Kalcsics and Ríos-Mercado, 2019), the majority of approaches include meta-heuristic or heuristic algorithms to handle the problem intractability. A major advantage of these methods is their flexibility to include almost any practical criterion and measure criteria design of districts. We develop a graph-based meta-heuristic algorithm based on an improved genetic algorithm (GA) to deal with the problem intractability and this algorithm aims in solving large-sized instances, efficiently.

There exists relatively rich literature on the *deterministic* districting problem. However, in many applications, effective districting decisions must be made in the presence of uncertainty. Stochastic programming models are proposed for the integrated districting and routing problems by Zhong et al. (2007), and Lei et al. (2016). Further, robust optimization (RO) models are developed for the graph partitioning involving interval uncertainty (Fan et al., 2012). In our healthcare districting problem, the demand for health services in each basic unit (i.e., residential area) has inherent uncertainty because of its seasonality and dependency on many parameters such as social culture, health services' price, patients' income, etc. (Harper et al., 2005; Feldstein, 1966). On the other hand, many issues such as population aging, the prevalence of chronic diseases, and the search for a higher quality of life can increase the demand value. Therefore, it is difficult to obtain a precise approximation for the demand value, although it is very essential to properly plan for a healthcare system. In this study, we address Iran's healthcare system and present a function based on the Hedonic regression model for the demand estimation at the first level of services. To the best of our knowledge, we present a robust healthcare districting problem for the first time in which we consider a deviation for the estimated value of health services demand and so interval uncertainty sets are employed.

Briefly, our main contributions in this paper are: (1) development of an optimization approach for the healthcare districting based on a new MIP model in accordance with the graph-based nature of the problem; (2) the presentation of a function for the demand approximation and the extension of the model for the robust decision-making under interval uncertainty; (3) The development of an improved graph-based GA for solving the optimization problem; (4) The investigation of a real-world case study based on data from Iran and the development of a framework for the hierarchical districting to address the districting at various levels.

The rest of this paper is organized as follows. In Section 2, we present a literature survey on the districting problem. In Section 3, the characteristics of our optimization problem are explained, and the deterministic and robust models are proposed. The estimation model of the healthcare services demnad is presented in Section 4. The developed solution algorithm is presented in Section 5. In Section 6, we present experimental results and sensitivity analyses for a real-world case study and several generated instances. Further, the driven managerial insights are briefly explained. Finally, Section 7 contains conclusions and future research directions.

2. Literature review

This section is presented in three parts. In the first part, the healthcare districting literature is reviewed. Districting problems in terms of optimization aspects are briefly discussed in the second part. Finally literature gaps are presented in the third part.

Healthcare districting problem. Recent advances in the field of districting are highlighted by an available book from Ríos-Mercado (2020) and in one chapter of this book, Yanık and Bozkaya (2020) reviewed the districting literature in the area of healthcare. Although there exists a rich literature on the general districting problem, a few papers examined healthcare districting. The pioneering study of Pezzella et al. (1981)) present a two-stage approach for partitioning a region into districts to ensure an optimal allocation of the available health services. Datta et al. (2013) propose a multi-objective optimization problem to geographically partition a territory into entities. They develop a multi-objective GA to solve their problem and examine a healthcare geography in the east of England as a real-world case study. Steiner et al. (2015) present a districting problem to improve the existing healthcare system in Parana State in Brazil with three objectives: maximization of the population homogeneity in the districts, maximization of the variety of medical procedures offered in the districts, and minimization of the inter-district distances to be traveled by patients. They adopt a multi-objective GA for solving the problem. Yanık et al. (2019) address a multi-period multi-criteria healthcare districting problem for the planning of primary healthcare services. In their study, an MIP model is proposed, which six criteria for the districts are taken into account. They examined their MIP-based approach by a case study in Istanbul, Turkey.

It should be mentioned that the home-care districting problem, in which home-care visits should be supplied for their inhabitants by using healthcare personnel such as nurses or physiotherapists, is studied by several studies (see e.g., Benzarti et al., 2013; Minciardi et al., 1981; Blais et al., 2003; Cortés et al., 2018; Gutiérrez-Gutiérrez and Vidal, 2015). The main requirements for the districts are equal travel time or workload, well accessibility, and the high utilization of the social facility's capacity. Recently, Lin et al. (in press) address the integrated elderly care service districting problem. In their study, a multi-objective mixed-integer nonlinear programming model is proposed, and a non-dominated sorting genetic algorithm II (NSGA II) is developed for solving their optimization problem.

Districting problems in terms of optimization aspects. In the general context of the districting problem, several mathematical models are proposed. Hess et al. (1965) presented the first formulation as an MIP model, and their model did not explicitly consider districts' contiguity. Then, Drexl and Haase (1999) embedded contiguity constraints into their optimization model by using a given neighborhood relation. As mentioned before, enforcing contiguity constraints by using a given neighborhood relation requires a number of inequalities that increase exponentially with the number of basic units, and make the model intractable for large problems (King et al., 2012). To deal with this challenge, for a commercial territory design planning, Ríos-Mercado and López-Pérez (2013) present a solution approach based on the cutting plane method and the structure of their problem that is able to solve large instances with small optimality gaps.

Furthermore, to model the contiguity constraints, based on Duque et al. (2011), there are three types of graph-based modeling approaches in the literature as 1- the geographical area is modeled as a forest of trees and each district should be a distinct tree in which the cycles are prevented by a set of constraints (see e.g., Yamada, 2009), 2- the basic units should be added to a district in a specified order, where orders prevent cycles and ensure contiguity (see e.g., Cova and Church, 2000), 3- by using a flow approach, (Shirabe, 2005, 2009) add continuous decision variables measuring the flow volume, and the districts' contiguity is guaranteed by establishing a unit flow from each basic unit within a district to a selected sink of the district. Based on empirical results of Duque et al. (2011), although a clear dominant model does not exist, the flow approach performs relatively better than other ones. However, Shirabe's model is also computationally intractable for large-sized problems. As other approaches for obtaining contiguous districts, Zoltners and Sinha (1983) use the shortest path adjacency tree to guarantee the contiguity of districts in the area of sales territory alignment, and the spanning tree-based approach is also developed by Kim (2018) for the political districting. Finally, it is worth noting that the most of above-mentioned approaches are proposed for the center-based districting problems in which one or many basic units should be selected for each district as its center(s) and its contiguity is ensured when there exists at least one path from other basic units to the corresponding center(s) without crossing the district's boundary. Such approaches are not applicable in our case and we develop new approaches to ensure contiguity of districts.

Various balancing criteria are also defined for the districting problem. Several studies (see e.g., Steiner et al., 2015; Datta et al., 2013; Salazar-Aguilar et al., 2013) assume that the size of each district has to be less than a predetermined upper bound and/or more than a given lower bound to obtain balanced districts. Further, the equitable workload and travel time are popular as the balancing criteria for the service districting problem.

Another well-known measure in the context of districting problems is the compactness of districts. A geographically compact district has a regular and round shape without distortion and holes. In the related literature, there are different variations of the proposed model by Hess et al. (1965) because of various compactness measures that are used in their objective function (See e.g., Salazar-Aguilar et al. (2011); Ríos-Mercado and Fernández (2009)). Unfortunately, as highlighted by Kalcsics and Ríos-Mercado (2019), there is not a general mathematical model for districting problems because of significant ambiguity on how to consider topological constraints and quantify the various planning criteria, such as balance and compactness, concurrently.

From an algorithmic perspective, to solve districting problems, several heuristics based on the location-allocation procedure (e.g., Hess et al., 1965; Ríos-Mercado et al., 2021), set partitioning approach (e.g., Garfinkel and Nemhauser, 1970), and seed-growing approach (e.g., Lei et al., 2012, 2015) are developed. Classical *meta*-heuristic algorithms such as Tabu Search (e.g., Ricca et al., 2013; Bozkaya et al., 2003), GRASP (e.g., Salazar-Aguilar et al., 2013; Ríos-Mercado and Fernández, 2009), Simulated Annealing (e.g., D'Amico et al., 2002), and GA (e.g., Steiner et al., 2015; Datta et al., 2012; Bacao et al., 2005) have been widely applied as well. Geometric approaches as another class of solution algorithms proposed by Kalcsics et al. (2005) and Galvão et al. (2006) utilize the districting problem's underlying geometrical information. Moreover, for the commercial territory design, Ríos-Mercado and López-Pérez (2013), Salazar-Aguilar et al. (2011), and Sandoval et al. (2020) propose an exact algorithm based on an iterative cut generation strategy within the branch and bound algorithm.

Although the methodological advances for modeling and solving districting problems are strongly motivated by complex real-world applications, there are a few papers in the literature dealing with districting decisions under uncertainty. For designing the last-mile delivery districts, a few studies such as (Zhong et al., 2007; Lei et al., 2012, 2016) developed two-stage stochastic programming models. Diglio et al. (2021) investigated a districting problem with stochastic demands and to deal with this problem, they developed chance-constrained balancing requirements.

In Table 1, several main studies in the area of the districting problem are categorized in terms of their key features including the model, solution approach, application, case study, and uncertainty consideration.

Literature gaps. From the above literature, it is highlighted that the districting problem in the area of healthcare operations management is scarcely addressed. The defined problem in this study is based on a real-world application in healthcare districting that is not tackled before in according to our best of knowledge. A significant gap needs to be addressed in order to improve and foster the body of knowledge in the districting problem is to investigate the impact of uncertainty on the districting problem by RO approaches. There is not any paper to address the interval uncertainty, and only a limited number of papers developed two-stage stochastic programs for the last-mile delivery districting. Furthermore, dynamic healthcare districting problem that can be suitable to deal with pandemics (e.g., outbreaks of COVID-19 in local communities) and new developments in OR, such as online optimization (Keyvanshokooh et al., 2021) and data-driven approaches (Keyvanshokooh et al., 2020; Fattahi and Govindan, 2020) are not addressed

Table 1
Literature survey on the districting problem.

Articles	Application	MIP model	Multiple objectives	Solution approach	Real case study	Input data
(D'Amico et al., 2002)	Police Districting			Simulated Annealing	✓	Certain
(Muyldermans et al., 2002)	Districting for salt spreading operations		✓	Heuristic	✓	Certain
(Blais et al., 2003)	Home healthcare districting			Tabu Search (TS)	✓	Certain
(Bozkaya et al., 2003)	Political districting			TS	✓	Certain
(Harper et al., 2005)	Healthcare districting			Geographical simulation	✓	Certain
(Haugland et al., 2007)	Last-mile delivery districting			TS		Uncertain
(Zhong et al., 2007)	Distribution districting		✓	TS		Uncertain
(Lin and Kao, 2009)	Municipal solid waste collection	✓			✓	Certain
(Ricca and Simeone, 2008)	Political Districting		✓	Local Search, TS, SA	✓	Certain
(Ríos-Mercado and Fernández, 2009)	Commercial territory design	✓	✓	GRASP		Certain
(Salazar-Aguilar et al., 2011)	Commercial territory design	✓		B&B and a cut generationstrategy		Certain
(Jarrah and Bard, 2012)	Districting in parcel delivery	✓		Column Generation	✓	Certain
(Datta et al., 2012)	Delineating census tracts		✓	GA	✓	Certain
(Lei et al., 2012)	Last-mile delivery districting			Large neighborhood search meta-heuristic	✓	Uncertain
(Salazar-Aguilar et al., 2013)	Commercial territory design		✓	NSGAII - GRASP		Certain
(Mayorga et al., 2013)	Emergency response planning			Simulation		Certain
(Datta et al., 2013)	Healthcare districting		✓	GA	✓	Certain
(Ríos-Mercado and López-Pérez, 2013)	Commercial territory design	✓		B&B and a cut generation strategy	✓	Certain
(Lei et al., 2015)	Last-mile delivery districting		✓	Adaptive large neighborhood search	✓	Certain
(Steiner et al., 2015)	Healthcare districting		✓	NSGAII	✓	Certain
(Konur and Geunes, 2019)	Distribution districting			Column Generation based heuristic		Certain
(Lei et al., 2016)	Last-mile delivery districting		✓	GA		Uncertain
(Yanık et al., 2019)	Healthcare districting	✓	✓			Certain
(Lin et al., in press)	Home healthcare districting		✓	NSGA II	✓	Certain
(Nasir et al., 2018)	Home healthcare districting	✓				Certain
(Sandoval et al., 2021)	Last-mile delivery districting	✓		A heuristic approach	✓	Certain
Our study	Healthcare Districting	✓		Graph-based GA	✓	Uncertain

in this area.

In summary, this study proposes an MIP model for the healthcare districting problem, which guarantees balanced and contiguous districts. Then, the extension of the model for the decision-making under interval uncertainty is developed. To handle the computational intractability of the problem, we propose a graph-based meta-heuristic algorithm. Finally, we investigate the applicability of the presented optimization tools for a real-world case study in Iran and provide some practical insights for healthcare practitioners.

3. Problem definition and formulation

This section is presented in four separate sub-sections as follow: in sub-section 3.1, our problem is described. A deterministic MIP model based on the spanning tree is proposed in sub-section 3.2. A heuristic method, entitled as enumeration-based approach, is presented in sub-section 3.3. Two RO approaches are extended in sub-section 3.4.

3.1. Problem definition

Here, we describe our healthcare districting problem along with its specific features. A geographical area, where a healthcare system operates, can be modeled as an undirected graph $G = (V, E)$ with a set of vertices $V = \{1, 2, ..., |V|\}$ as residential areas in which $|V|$ represents the cardinality of set V, and $E = \{e_{ij} | 1 \leqslant i, j \leqslant |V|, i \neq j,\ e_{ij} = e_{ji}\}$, is a set of edges where e_{ij} represents the edge between i and j as the traveling direct route. In each residential area, the demand for the healthcare services is described as the population size that requires health services in a specified time interval (See e.g., Harper et al., 2005; Blais et al., 2003). The capacity of the healthcare system and the population size requiring health services are denoted by δ_i and α_i for each residential area i, $i = 1, 2, ..., |V|$, respectively.

General conditions. Healthcare districting is defined as partitioning residential areas into a set of districts in order to improve the provision of healthcare services. A district D_p, $p = 1, 2, ..., |P|$, corresponds to a component of graph G, i.e., $D_p = (V_p, E_p) \subseteq G$ so that $V_p \subseteq V$, $E_p \subseteq E$, and $e_{ij} \in E_p$ with $i.j \in V_p$. As a consequence, the set of non-empty and disjoint partitions related to the residential areas is $\{V_1, V_2, ..., V_p, ..., V_{|P|}\}$. Some general conditions (mathematical expressions (1)) required for partitioning the residential areas are as follows:

$$\begin{array}{ll} 2 \leqslant |P| \leqslant |V| & \\ V_p \neq \varnothing & p = 1, 2, ..., |P| \\ V_p \cap V_{p'} = \varnothing & p, p' = 1, 2, ..., |P|, p \neq p' \\ V_1 \cup V_2 \cup ... \cup V_{|P|} = V & \end{array} \quad (1)$$

Objective function. In the healthcare districting problem, in various districts, the difference between the capacity and the demand related to the healthcare system should be the same as much as possible to achieve balanced districts. Therefore, criterion $\omega(D_p)$ is defined for district p as $\sum_{i \in V_p} (\delta_i - \alpha_i)$, and the objective function (mathematical expression (2)) is calculated as follows:

$$Min \left(\underset{p, p' \in P:\ p \neq p'}{Max} \{ \omega(D_p) - \omega(D_{p'}) \} \right) \quad (2)$$

Contiguity of districts. The residential areas corresponding to a district should be inter-connected. In other words, a district cannot be formed by several disconnected areas. To ensure the contiguity of districts, we define a graph-based condition. From a topological point of view, district D_p is said to be contiguous if it is possible to travel between each two different nodes of the district without crossing the district's boundary. Spanning tree T corresponds to graph G is an undirected subgraph of G in which any two vertices of G are connected by a unique path. In this study, a set $S_{ij}^{(T)}$ contains the vertices in the corresponding path from vertex i to j in tree T except i and j. As a consequence, we define the following sufficient condition to have contiguous district D_p.

Condition (1) *:* By assuming T as a spanning tree of undirected graph G, if two different vertices i and j are in V_p, then vertices in $S_{ij}^{(T)}$ should also be in V_p.

The maximum traveling time/distance through districts. An important issue in the healthcare districting problem is to limit the length of travel time through paths of a district. By addressing this issue, we can improve the ease of the health network management through a district, and the patients and health workers can easily travel within the district. Let parameter τ_{ij} be the weight of edge e_{ij} that shows the travelling route distance between vertices i and j. In this study, it is assumed that $\tau_{ij} = \tau_{ji}$, where $1\leqslant i, j\leqslant |V|$, and $i \neq j$. Then, $L_{ij}^{(T)}$ shows the length of corresponding path from vertex i to j in tree T that can be computed as the sum of weights related to edges of the path. Then, we consider the following condition to restrict the geographical extent of district D_p.

Condition (2)*:* By assuming T as a spanning tree of undirected graph G, if two different vertices i and j are in V_p, $L_{ij}^{(T)}$ should be less than or equal to a predetermined value L^{max}.

Based on Datta et al. (2013), another condition should be considered for the healthcare districting problem to limit the number of districts:

Condition (3)*:* The healthcare map should be partitioned into a number of districts ($|P|$) that is within a given range or at a fixed value, i.e., $P_{min}\leqslant|P|\leqslant P_{max}$ or $|P| = P_{\text{fix}}$.

3.2. Spanning tree-based MIP model

In this sub-section, we propose an MIP model corresponding to predetermined spanning tree T for the healthcare districting problem with the presented conditions. The used notations are presented in Table 2.

The deterministic formulation of the problem is:

$$Min: \ z \tag{3}$$

$$\sum_{i\in V}(\delta_i - \alpha_i)x_{ip} - \sum_{i\in V}(\delta_i - \alpha_i)x_{ip'} \leqslant z, \forall p, p' \in P : p \neq p', \tag{4}$$

$$\sum_{p\in P} x_{ip} = 1, \forall i \in V, \tag{5}$$

$$\sum_{k\in S_{ij}^{(T)}} x_{kp} \geqslant \left|S_{ij}^{(T)}\right| \left(x_{ip} + x_{jp} - 1\right), \forall p \in P, \forall i,j \in V, \tag{6}$$

$$x_{ip} + x_{jp} \leqslant 1, \forall i,j \in V : L_{ij}^{(T)} > L^{max}, \forall p \in P, \tag{7}$$

$$x_{ip} \in \{0,1\}, \forall p \in P, \forall i \in V. \tag{8}$$

Difference between the capacity and demand of the healthcare system in each district is defined as a criterion to represent the healthcare system ability in serving the corresponding patients. Objective function (3), according to constraints (4), minimizes the maximum difference of this criterion in various districts to achieve balanced districts. Constraints (5) guarantee that each node, basic unit, should be assigned to only one district. Constraints (6) and (7) ensure *conditions* (1) and (2) for the mathematical model, respectively. In accordance with *condition* (1), if two different nodes i and j belong to district p, other nodes in the path from node i to j in the considered tree T should also be in district p. Based on *condition* (2), in each district, the maximum length of each path corresponding to tree T should be less than or equal to an upper bound value, L^{max}. Finally, Constraints (8) specify binary decision variables.

Table 2
The notations.

Sets and Parameters	
V	Set of basic units, nodes, $i,j,k \in V$, $i,j,k = 1,2,\ldots,|V|$,
P	Set of districts,$p,p' \in P$, *and* $p,p' = 1,2,\ldots,|P|$,
$S_{ij}^{(T)}$	Set of vertices in the corresponding path from node i to j in predetermined spanning tree T except i and j, and $\left|S_{ij}^{(T)}\right|$ represents the cardinality of set $S_{ij}^{(T)}$.
$L_{ij}^{(T)}$	The length of the path from node i to j in tree T,
L^{max}	The maximum allowable value for the length of a path between two different nodes in a district,
α_i	The number of patients requiring health services in node i,
δ_i	The capacity of the healthcare system in node i for the service provision,
Decision variables:	
x_{ip}	A binary variable that equals 1 if node i is assigned to district p, and 0 otherwise.
z	A real-valued variable representing the value of the problem's objective function.

3.3. Enumeration-based approach

The solution of the proposed MIP model is dependent on the considered spanning tree T. To find a good solution for the healthcare districting problem, we solve the MIP model for the all shortest-path trees corresponding to graph G. For a connected graph G, a shortest-path tree rooted at vertex $i \in V$ is a spanning tree of G such that the path's length from vertex i to other vertices in the tree has the minimum length among existing paths in G. For example, Fig. 1 shows an illustration for constructing a shortest-path tree rooted at vertex 4 in a connected graph with 9 vertices and 15 edges. In Fig. 1, the weight of the edges is assumed to be the traveling route distance.

The proposed enumeration-based approach is explained in Algorithm I.

Algorithm I: *The enumeration-based algorithm*

Input: Connected graph G related to a geographical area,$G = (V, A)$
For each $i \in V$
  Construct the shortest-path tree $T^S(i)$ rooted at vertex i;
  Solve the spanning tree-based MIP model related to $T^S(i)$;
  Save the optimal districting solution $x^*(i)$ and objective function $z^*(i)$;
End
Return the final objective function z_E^* as $z_E^* = \min_{i\in V}\{z^*(i)\}$.
Return the corresponding best districting solution (x_E^*).

In the enumeration-based algorithm, to find a shortest-path between two various vertices in a connected graph, *Dijkstra* algorithm with time complexity $O\left(|V|^2\right)$ is used. In Algorithm I, the number of final solutions is equal to the number of nodes so that the optimal objective of the MIP model related to their corresponding shortest-path tree is z_E^*. In other words, the algorithm would have a unique solution, if z_E^* is the optimal objective value of the MIP model related to a unique shortest-path tree of graph G. In the Supplementary materials, we modify the proposed MIP model to present an alternative MIP model, entitled as shortest-path-based model, in which a district is considered contiguous if its basic units are connected by shortest paths.

3.4. RO models for uncertain demands

In this study, in each residential area, the nominal value of the demand for health services is obtained as a percentage of the population size. In the presented optimization problem, α_i, $i = 1,2,\ldots,|V|$, represents the size of population in each node that requires healthcare services. We address the interval uncertainty for weight vector $\boldsymbol{\alpha} = (\alpha_i)_{|V|\times 1}$. Each entry α_i is modeled as independent, symmetric, and bounded but unknown distribution variable $\widetilde{\alpha}_i$ that takes value in $I_i = \left[\overline{\alpha}_i - \widehat{\alpha}_i, \overline{\alpha}_i + \widehat{\alpha}_i\right]$ where $\overline{\alpha}_i\geqslant 0$ is the nominal value, $\widehat{\alpha}_i\geqslant 0$ is the maximum deviation from this nominal value, and $\overline{\alpha}_i\geqslant\widehat{\alpha}_i$. The robust version of constraints (4) can be formulated as:

$$\max_{\forall \widetilde{\alpha}_i \in I_i} \left(\sum_{i \in V} \left(\delta_i - \widetilde{\alpha}_i \right) x_{ip} - \sum_{i \in V} \left(\delta_i - \widetilde{\alpha}_i \right) x_{ip'} \right) \leqslant z, \ \forall p, p' \in P : p \neq p'. \tag{9}$$

In this paper, we will present two RO models and examine them in the computational results.

Bertsimas & Sim RO approach. The RO approach presented by Bertsimas and Sim (Bertsimas and Sim, 2004) adjusts the level of the conservatism of robust solutions in terms of probabilistic bounds of constraint violations. This formulation is tractable and hence this approach has become popular, recently. To reformulate the problem based on this method, by assuming $\widetilde{\Delta}_i = \delta_i - \widetilde{\alpha}_i$, where $\widetilde{\Delta}_i \in I_i^{\Delta} = \left[\overline{\Delta}_i - \widehat{\Delta}_i, \overline{\Delta}_i + \widehat{\Delta}_i \right]$, constraints (9) can be written as follows:

$$\max_{\forall \widetilde{\Delta}_i \in I_i^{\Delta}} \left(\sum_{i \in V} \widetilde{\Delta}_i x_{ip} - \sum_{i \in V} \widetilde{\Delta}_i x_{ip'} \right) \leqslant z, \forall p, p' \in P : p \neq p'. \tag{10}$$

It is unlikely that all uncertain parameters will realize their worst-case values simultaneously. Thus, a maximum number of parameters that can deviate from their nominal values should be limited. We define scaled deviations $k_{ip} = \frac{\widetilde{\Delta}_i - \overline{\Delta}_i}{\widehat{\Delta}_i}$ that always belongs to the interval $[-1, +1]$ and the aggregated scaled deviation of uncertain parameters are limited to Γ for each $p \in P$ as $\sum_{i \in V} |k_{ip}| \leqslant \Gamma$. Γ is a parameter, not necessarily integer, that takes values in the interval $[0, |V|]$. Therefore, $I_i^{\Delta} = \left\{ \widetilde{\Delta}_i \middle| \widetilde{\Delta}_i = \widehat{\Delta}_i k_{ip} + \overline{\Delta}_i, \forall \boldsymbol{k}_p \in \Omega_p \right\}$ and $\Omega_p = \left\{ \boldsymbol{k}_p \middle| \sum_{i \in V} |k_{ip}| \leqslant \Gamma, |k_{ip}| \leqslant 1, \ \forall i \in V \right\}$. As a consequence, constraints (10) is equal to:

$$\begin{cases} \min z \\ s.t. \sum_{i \in V} \overline{\Delta}_i x_{ip} - \sum_{i \in V} \overline{\Delta}_i x_{ip'} + \max_{k_p \in \Omega_p, k_{p'} \in \Omega_{p'}} \left\{ \sum_{i \in V} \widehat{\Delta}_i x_{ip} k_{ip} - \sum_{i \in V} \widehat{\Delta}_i x_{ip'} k_{ip'} \right\} \leqslant z, \forall p, p' \in P \end{cases} \tag{11}$$

For a given value of $\boldsymbol{x}^*$, the scaled deviations k_{ip} will take positive and $k_{ip'}$ will take negative value in optimization problem (11). As a consequence, the inner maximization in optimization problem (11) is equal to:

$$\begin{cases} \max_{p, p' \in P: p \neq p'} \left\{ \sum_{i \in V} \widehat{\Delta}_i x_{ip}^* k_{ip} + \sum_{i \in V} \widehat{\Delta}_i x_{ip'}^* k_{ip'} \right\} \\ \sum_{i \in V} k_{ip} \leqslant \Gamma, \quad \forall p \in P \\ 0 \leqslant k_{ip} \leqslant 1, \quad \forall p \in P, \forall i \in V, \end{cases} \tag{12}$$

To obtain a tractable optimization from (12), we can dualize the inner maximization after introducing dual variables u_p and w_{ip}. For more information regarding this approach, one can refer to Bertsimas and Sim (Bertsimas and Sim, 2004). Therefore, we can obtain the following optimization problem (13) that is equal to constraints (10).

$$\begin{cases} \min z \\ s.t. \ \sum_{i \in V} \overline{\Delta}_i x_{ip} - \sum_{i \in V} \overline{\Delta}_i x_{ip'} + \Gamma\left(u_p + u_{p'}\right) + \sum_{i \in V} w_{ip} + \sum_{i \in V} w_{ip'} \leqslant z \ , & \forall p, p' \in P, \\ w_{ip} + u_p \geqslant \widehat{\Delta}_i x_{ip} \ , & \forall p \in P, \forall i \in V, \\ u_p \geqslant 0, \ w_{ip} \geqslant 0, & \forall p \in P, \forall i \in V, \end{cases} \tag{13}$$

By combining optimization problem (13) with our problem, we can obtain a tractable MIP formulation, called BSRO. If $\Gamma = 0$, it reduces to the nominal formulation where there is no protection against uncertainty . If $\Gamma = |V|$, the problem is completely protected against the worst-case realization of uncertain parameters.

Weighted protestion against worst-case condition. In this paper, we also develop a simple RO approach for dealing with the demand uncertainty as *weighted protection against worst-case condition*, called as WWRO. For the highest protection against uncertainty, we consider the worst-case condition in which the maximum violations are taken into account. Since $\sum_{p \in P} x_{ip} = 1, \forall i \in V$, constraints (4) can be adopted for the worst-case condition as following form:

$$\sum_{i \in V} \delta_i x_{ip} - \sum_{i \in V} \delta_i x_{ip'} - \sum_{i \in V} \left(\overline{\alpha}_i - \widehat{\alpha}_i \right) x_{ip} + \sum_{i \in V} \left(\overline{\alpha}_i + \widehat{\alpha}_i \right) x_{ip'} \leqslant z, \forall p, p' \in P : p \neq p'. \tag{14}$$

To simply adjust the solution's level of conservatism against the robustness, we can define a control parameter $0 \leqslant \lambda \leqslant 1$, and write constraints (14) in the robust counterpart formulation as follows:

$$\sum_{i \in V} \left(\delta_i - \overline{\alpha}_i \right) x_{ip} - \sum_{i \in V} \left(\delta_i - \overline{\alpha}_i \right) x_{ip'} + \lambda \left(\sum_{i \in V} \widehat{\alpha}_i \left(x_{ip} + x_{ip'} \right) \right) \leqslant z, \forall p, p' \in P : p \neq p'. \tag{15}$$

Therefore, if $\lambda = 0$, constraints (15) reduce to the nominal formulation where there is no protection against uncertainty, and if $\lambda = 1$, constraints are fully protected against the worst-case realization of the uncertainy.

4. Demand estimation of healthcare services

Obtaining the demand value per basic unit is a challenging issue, although, there are several studies in the literature addressing this issue by econometric/micro-econometric models (see e.g., Akin et al., 1995; de Sousa and Gomes, 2020; Jack, 1999). One of the popular methods for this purpose is to use Hedonic models, applied by Jack (1999), Goldman and Grossman (1978), and Gertler and Gaag (1990). Hedonic demand theory is a revealed preference method of estimating the demand of a good or service, which should be broken down into essential characteristics and obtained by estimating the contributory value of each characteristic. Hedonic models are most commonly estimated by using regression analysis (Sopranzetti, 2015).

Forecast of healthcare demand for each service level will need to account population demographics. Based on the literature, including (Harper et al., 2005; Yanık et al., 2019; Feldstein, 1966; and Karimi, 2019), and the analysis of our case study, *the population aging*, *the percentage of married women*, *the unemployment rate*, *the average of income level*, *the literacy level*, *the urban development*, *the percentage of men*, *the population density*, and *the environmental factors* such as air pollution in compared with their standard level are considered as main population

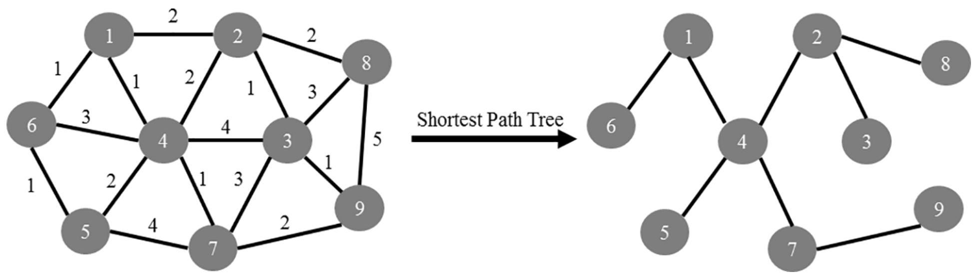


Fig. 1. The illustration of a shortest-path tree rooted at vertex 4 in a connected graph with 9 vertices and 15 edges.

attributes, independent variables, affecting the healthcare demand.

Three main transform structures are popular in Hedonic models consist of *linear*, *semi*-log, and *Box-Cox*. We use Box-Cox in this study because of its generality that is presented in relation (16) and the regression model is based on relation (17).

$$X^{\lambda}=\begin{cases}\dfrac{X^{\lambda}-1}{\lambda} & \lambda\neq 0, X>0\\ \ln X, & \lambda=0\end{cases} \tag{16}$$

$$D=X^{\lambda}\beta+\varepsilon \tag{17}$$

In relations (16) and (17), X is the independent variables vector, β is the variables' coefficient, λ is Box-Cox transform parameter, D is the demand estimation (the population percentage that requires health services), and ε is the error term following normal distribution function with mean zero and standard deviation σ. Based on (Sopranzetti, 2015), we have used the log-likelihood function (18) to obtain a good and tractable approximation for model's parameters, where y_i is the real value of demand in i^{th} sample of available data, z represents the standard normal variable, and L is the likelihood function. In this relation, we have the log of the likelihood function, the values of the parameters $(\beta, \sigma^2, \lambda)$are determined by maximum likelihood estimation. It is worth noting that D is the population percentage that requires health services and the demand value can be obtained by multiplying D by the number of population.

$$ln\,L(\beta,\sigma^2,\lambda|y)=\frac{-1}{2\sigma^2}(z-X\beta)(z-X\beta)-\frac{n}{2}ln(2\pi\sigma^2)+(\lambda-1)\sum_{i=1}^{n}ln(y_i) \tag{18}$$

5. Solution algorithm

Due to the problem's NP-hardness (Kalcsics and Ríos-Mercado, 2019), we develop GA-based meta-heuristic approaches to deal with the computational intractability of the optimization problem which is able to provide near-optimal solution in a reasonable run time. The algorithms are based on the graph-based nature of the problem. GA and its improved version are presented in sub-sections 5–1 and 5–2, respectively. Finally, in sub-section 5–3, the parameter setting of algorithms are explained. It should be mentioned that we have examined several meta-heuristic algorithms and the proposed one in the paper had an acceptable performance for our problem.

5.1. Genetic algorithm

The GA is an evolutionary meta-heuristic algorithm, which is inspired by the process of natural selection (Mitchell, 1998), and has been applied for solving the districting problem by several studies. Generally, in a GA, a solution for the optimization problem should be coded as an array, called a chromosome, and a set of solutions, called a population (generation), evolves toward better solutions. Each solution of the population has an associated fitness which is calculated based on the problem's objective function. The evolution in the GA is an iterative process that typically starts from a population of randomly generated chromosomes, and the chromosomes with better fitness value are more likely to survive to the next generation and consequently, the algorithm converges to an optimal or near-optimal solution. To obtain the next generation in each GA's iteration, bio-inspired operators such as mutation, crossover, and selection are commonly used. In this study, after initializing solutions, we apply the tournament selection operator in which we select a random sub-group from the population and then the solutions with the highest fitness are selected to create a mating pool. Genetic operators, including crossover and mutation, are then applied successively to the mating pool to create a new set of solutions (offspring). The steps of the GA is shown in Fig. A1 in Appendix A. In Fig. A1 and the GA explanation, the size of population and mating pool are illustrated by pop_size and mpool_size, respectively. Further, the crossover and mutation probability are shown by crossover_prob and mutation_prob, successively.

Chromosome representation. It is difficult to obtain a feasible solution for the districting problem through random assignment (Steiner et al., 2015). Therefore, a new approach based on the shortest-path tree is applied to generate a feasible solution in terms of *condition* (1) without repairing mechanism. Each solution for partitioning graph $G(V,E)$ into $|P|$ sub-graphs can be obtained by eliminating $|P|-1$ edges of a tree with $|V|-1$ edges related to graph G. Thus, shortest-path tree $T^S(i)$ for each $i\in V$ in graph G can represent a solution for the problem after eliminating $|P|-1$ edges. In the obtained solution by this approach, *condition* (1) is guaranteed and the obtained districts are connected. In the proposed GA, an array of $|P|$ elements represents a problem's solution. The first element, i, has an integer value between 1 and $|V|$ that presents shortest-path tree $T^S(i)$, and other elements have different integer values between 1 and $|V|-1$ that represent eliminated edges in the corresponding shortest-path tree. In Fig. 2, the obtained solution for the districting problem is represented for a GA's solution.

In each solution of the GA, only the removed edges of a specified tree are shown. However, the set of vertices in each district is required to calculate the fitness function of the solution, and hence Breadth-First Search (BFS) algorithm is used to specify the vertices in each district. BFS is a well-known algorithm for traversing or searching tree or graph data structures, and its application in finding connected components of graphs are explained by Farahani (Farahani, 2012). The time complexity of the BFS algorithm for finding the connected components in a graph $G(V,E)$ is $O(|V|+|E|)$ for the worst-case performance when every vertex and edge is explored.

The main advantage of our proposed solution scheme for the GA is the fact that it ensures *condition* (1) and the connectivity of the obtained districts. In each population of the GA, the solutions that violate *condition* (2) are eliminated. The fitness function value (Eq. (19)) for each solution $\tilde{x}$ based on the WWRO method, $o(\tilde{x})$, is calculated as:

$$o(\tilde{x})=\max_{p,p'\in P:p\neq p'}\left\{\sum_{i\in V}\left(\delta_i-\overline{\alpha}_i\right)\tilde{x}_{ip}-\sum_{i\in V}\left(\delta_i-\overline{\alpha}_i\right)\tilde{x}_{ip'}+\lambda\left(\sum_{i\in V}\widehat{\alpha}_i\left(\tilde{x}_{ip}+\tilde{x}_{ip'}\right)\right)\right\}, \tag{19}$$

As mentioned before, to obtain the next generation, the mating pool

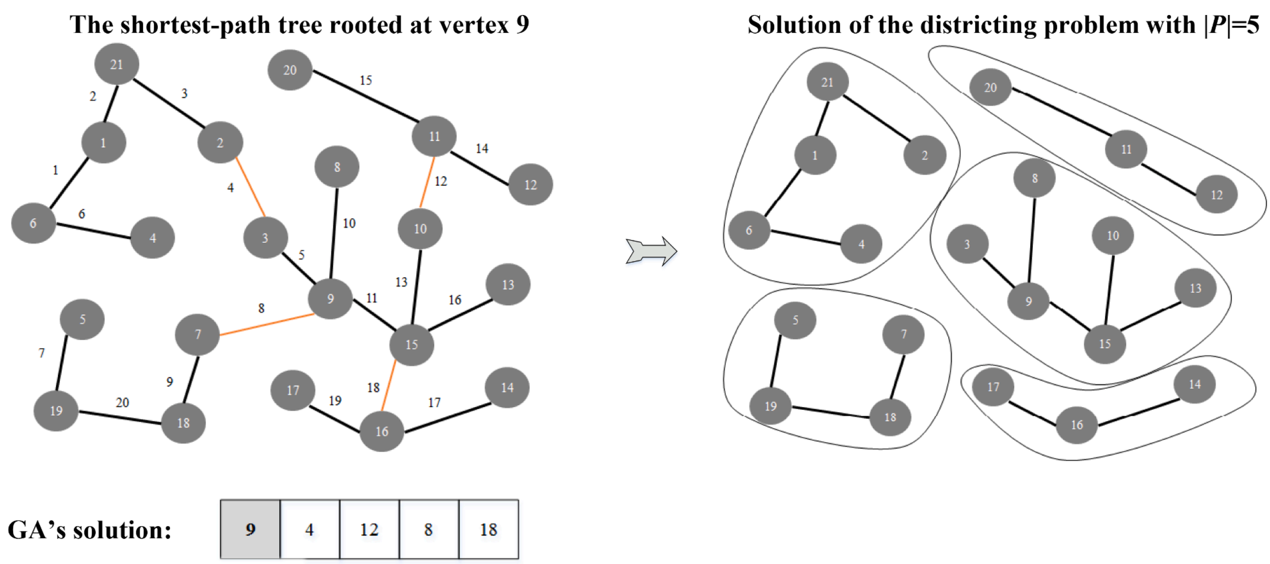


Fig. 2. An illustrative example for encoding and decoding procedure of GA's solutions.

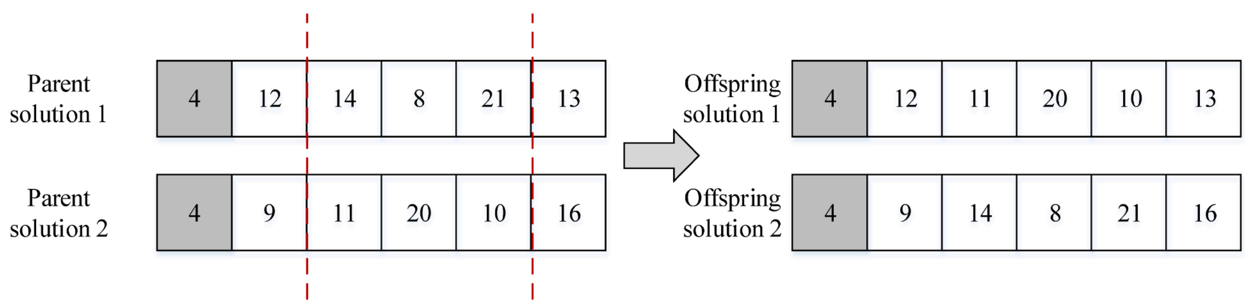


Fig. 3. The illustration of the crossover operator.

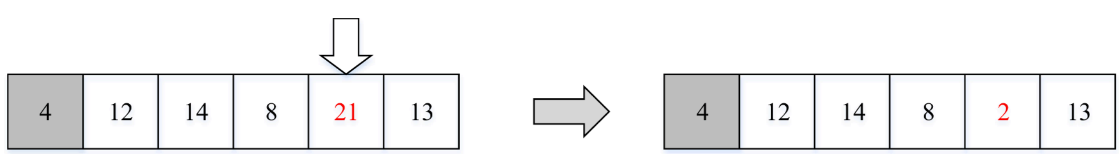


Fig. 4. The illustration of the mutation operator.

is constructed by the selected individuals, and the mutation and crossover operators, are successively implemented on them with predetermined probabilities to generate offspring.

Crossover operator. This operator is a two-point crossover with repairment of infeasible solutions. For the repairment mechanism, during the crossover operator, we take care of duplicating a value corresponding to the eliminated edges in new solution arrays. In such a situation, duplicated values are changed to the nearest possible value. Fig. 3 provides an example of the crossover operator.

Mutation operator. This operator randomly selects an element in the solution array of an offspring generated by the crossover operator, and changes its value. During the mutation operator, it takes care of duplicating a value corresponding to eliminated edges in a new solution array. Fig. 4 illustrates the mutation operator.

Preprocessing stage. We define the elementary basic units composing the territory for a healthcare system operation as an undirected and connected graph $G(V,E)$. In graph G where shortest-paths are well-defined, at the preprocessing stage, we construct shortest-path tree $T^S(i)$ for each $i \in V$. To obtain shortest-path trees, Dijkstra algorithm is used to find the shortest-path from a source vertex to a destination vertex (Wu and Chao, 2006).

Further, the edges in graph G are labeled from 1 to $|E|$, and the edges of each shortest-path tree $T^S(i)$ are also labeled from 1 to $|V|-1$. In shortest-path trees, the edge numbers are based on their sequence in graph G. Moreover, the length of the path between each two different vertices in shortest-path trees is calculated.

5.2. Improved genetic algorithm

Based on this sub-section, we improve the GA by two well-known mechanisms.

Mating pool construction with remainder stochastic sampling without replacement policy. To examine another version of the GA, the Remainder Stochastic Sampling without Replacement (RSSWR) policy is applied for constructing the mating pool as a well-known approach (Wicks and Reasor, 1999). This approach reduces the stochastic errors associated with the traditional selection approaches. In this method, the selection probability of individual q in the mating pool is obtained based on its fitness function value as $Pr_q = \frac{\left(1/o\left(\widetilde{x}_q\right)\right)}{\sum_{q\in Q}\left(1/o\left(\widetilde{x}_q\right)\right)}$, where Q is the set of all individuals in the population. As a consequence, the expected number of copies for each individual q, $e(q)$, is calculated as $e(q) = Pr_q \times$ mpoolsize.

The integer part of $e(q)$ is shown by $I(q)$, and in the RSSWR policy, individual q is deterministically copied $I(q)$ times in the mating pool. To fill out the rest of the mating pool, the fractional parts related to the expected number of copies of the individuals are normalized to the sum of one and assumed as probabilities for the selection process. As a consequence, for each individual, a uniform number between zero and one is randomly generated and if the generated number is less than or equal to its corresponding probability, an additional copy of the individual adds into the mating pool, and the probability of adding another copy of the individual is set to zero. This process continues, by adding at most one additional copy of each population's individual, until the number of individuals in the mating pool becomes equal to its desired size.

Local search procedure for the GA. Several studies use local search techniques to improve a meta-heuristic algorithm (See e.g., Calvete et al., 2016; Eroglu and Kilic, 2017). The GA is also improved by incorporating a local search procedure. In each iteration of the algorithm, after implementing the GA operators, the local search is performed on the obtained new generation and the best individuals are selected to form the population of the next iteration. To make new solution by using the local search procedure, each element of a solution array is added by a random number from the normal distribution with mean 0 and standard deviation 1, and then rounded to the nearest integer number. Moreover, we prevent form duplicating a value corresponding to the eliminated edges in a solution array and obtaining values out of their feasible bound. It should be noted that the mean and standard deviation values for the normal distribution had a good performance for the IGA in a large number of experiments using several instances. Fig. 5 illustrates the local search procedure for a solution array.

At each iteration of the algorithm, the local search procedure is implemented for each individual of the population two times and only the improved solutions are accepted. Although the implementation of the local searching within the GA is time consuming, it can enhance the performance of the algorithm for our optimization problem in terms of the objective function, meaningfully. The GA with the RSSWR policy and local search procedure is denoted by IGA in the remainder of the paper.

5.3. Setting the solution algorithms' parameters

In addition, the performance of meta-heuristic algorithms is considerably sensitive to values of their user-defined parameters because of the stochastic nature of them. In order to set these parameters, called controllable factors, for these algorithms to be suitable for solving problem instances with different sizes, Taguchi experimental design is applied and the objective function of the algorithms is considered as the response variable of the design. Taguchi method as a statistical method uses much fewer experiments in comparison with the full fractional design as well as minimizing variation of output results (see Montgomery, 2001; Taguchi, 1986). For more explanations about this method, one can refer to Fattahi et al. (2014); and Zarandi et al. (2013). Table 3 presents the final selected levels of factors for the GA and IGA in which 10,000 iterations are considered for performing them.

Table 3
Setting of GA's parameters.

	pop_size	mpool_size	crossover_prob	mutation_prob
GA	$2.5 \times \|V\|$	$0.8 \times$ pop_size	0.9	0.1
IGA	$2.5 \times \|V\|$	$1 \times$ pop_size	0.9	0.05

6. Computational results

Experimental results are presented for Iran healthcare system as a real-world case study and several randomly generated problem instances. In this section, the applicability of the mathematical models and the performance of the developed solution algorithms are investigated. Further, by conducting sensitivity analyses, some managerial insights are driven. In this study, the proposed models are solved by commercial software CPLEX 12 and the solution algorithms are coded in the C# programming language. A personal computer with Intel Core i7-640 M CPU (3.2 GHz), with 16 GB of RAM, is used for all implementations. We generate 40 problem instances (P1-P40) in three different data sets (S1, S2, and S3) in which P1-P20 are small- or medium-sized that are solvable by the CPLEX, and P21-P40 are medium or large-sized problem instances. The method of generation of problem instances is explained in Appendix B.

We solve problem instances P1-P20, including S1, S2, and S3 types, by the enumeration-based approach. The enumeration-based approach does not consider all spanning trees corresponding to a geographical area's connected graph and the spanning tree-based MIP model is solved for all shortest-path trees. As a consequence, the enumeration-based approach is an approximation. It worth noting that the final objective of the enumeration-based approach will be utilized to investigate the efficiency of the GA algorithms. We also compare its solution with the solution of a mathematical model that is proposed based on (Shirabe, 2005, 2009). This model, entitled Shirabe's model, is presented in the Supplementary Materials. In Shirabe's model, the contiguous districts are guaranteed by using a flow approach in which continuous decision variables are added to measure the flow volume. Shirabe's model considers the all feasible contiguous districts. The details of our computational results related to the above-mentioned problem instances are also presented in the Supplementary Materials. The corresponding results show that although all considered problem instances are solvable by the enumeration-based approach, the optimal objective of our problem is obtained by solving Shirabe's model in some small-sized instances P1-P8 and the CPLEX has run out of memory for other instances. Further, the average optimality gap of our enumeration-based approach is 2.01% in these instances.

6.1. Results of solving problem instances using solution algorithms

In this sub-section, several problem instances with different sizes are solved by using the GA and IGA, and the control parameter λ is

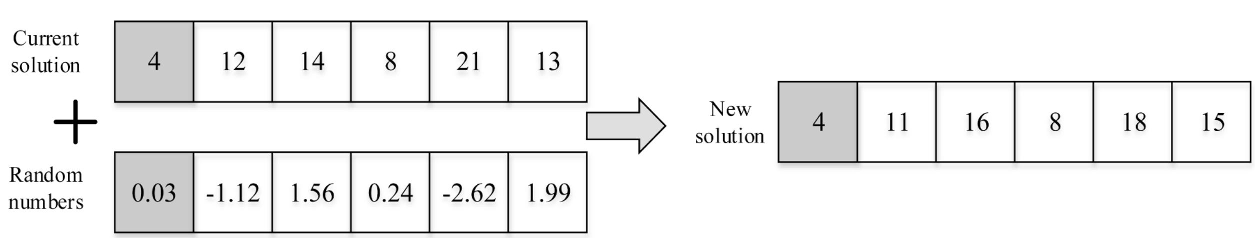


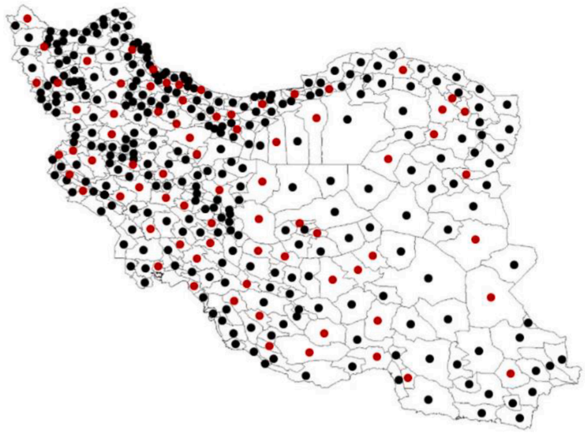
Fig. 5. the local search procedure for a solution array.

Table 4
Examination of the solution algorithms with the optimal solution of the problem.

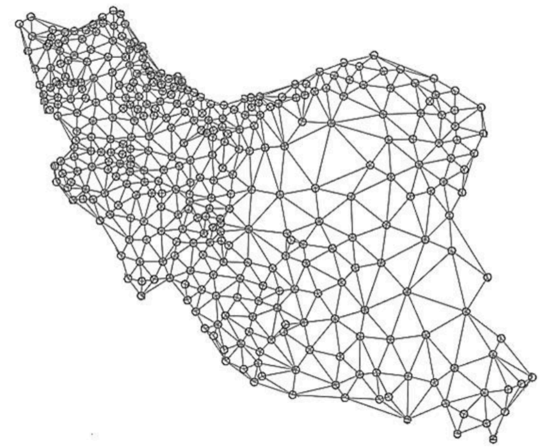
Tests	z_E^*	GA				IGA			
		Objective value	CPU time (S)	E-gap (%)	S-gap (%)	Objective value	CPU time (S)	E-gap (%)	S-gap (%)
P1-S1	6471	6549	78	1.21	3.50	6471	129	0	2.35
P1-S2	5311	5374	78	1.18	3.36	5311	135	0	2.15
P1-S3	6638	6638	75	0	2.32	6638	133	0	2.32
P2-S1	8916	9036	94	1.35	2.98	8918	139	0.02	1.64
P2-S2	14,011	14,186	90	1.25	2.12	14,011	140	0	0.86
P2-S3	12,737	12,892	97	1.22	3.18	12,737	165	0	1.94
P3-S1	14,910	15,344	109	2.91	6.17	14,943	191	0.21	3.39
P3-S2	19,880	20,455	117	2.89	4.72	19,884	189	0.02	1.80
P3-S3	24,850	25,491	122	2.58	4.75	24,882	216	0.13	2.24
P4-S1	18,473	19,147	125	3.65	5.78	18,503	245	0.16	2.22
P4-S2	13,855	14,363	138	3.67	6.14	13,876	275	0.15	2.54
P4-S3	15,394	15,853	152	2.98	6.39	15,466	269	0.47	3.79
P5-S1	7784	8018	154	3.01	4.34	7789	271	0.06	1.36
P5-S2	9014	9313	166	3.32	4.34	9055	308	0.45	1.44
P5-S3	11,475	11,869	177	3.43	6.84	11,501	318	0.22	3.52
P6-S1	14,768	15,328	165	3.79	6.86	14,878	276	0.01	3.73
P6-S2	16,408	17,025	168	3.76	6.34	16,639	277	1.41	3.93
P6-S3	14,502	15,040	177	3.71	7.09	14,595	335	0.64	3.92
P7-S1	7955	8127	189	2.16	2.74	7990	280	0.43	1.01
P7-S2	20,754	21,165	198	1.98	4.28	20,766	283	0.06	2.32
P7-S3	31,774	32,028	194	0.8	3.20	31,949	340	0.55	2.94
P8-S1	111,463	112,945	194	1.33	2.75	111,809	334	0.31	1.72
P8-S2	95,540	98,014	200	2.59	3.64	96,151	298	0.64	1.67
P8-S3	79,616	81,383	202	2.22	3.49	80,980	322	1.71	2.98
P9-S1	79,885	81,786	209	2.38	—	81,108	351	1.53	—
P9-S2	106,514	111,275	232	4.47	—	107,345	343	0.78	—
P9-S3	88,761	90,820	231	2.32	—	88,826	403	0.07	—
P10-S1	71,414	72,571	247	1.62	—	72,356	350	1.32	—
P10-S2	95,217	99,130	241	4.11	—	95,407	432	0.2	—
P10-S3	119,008	122,043	259	2.55	—	120,551	416	1.3	—
P11-S1	150,574	153,616	260	2.02	—	151,487	468	0.61	—
P11-S2	135,517	140,883	245	3.96	—	138,011	383	1.84	—
P11-S3	151,004	154,190	273	2.11	—	151,715	522	0.47	—
P12-S1	87,106	90,163	273	3.51	—	87,124	432	0.02	—
P12-S2	136,881	142,082	264	3.8	—	138,606	468	1.26	—
P12-S3	124,437	128,319	273	3.12	—	125,992	561	1.25	—
P13-S1	134,627	138,639	294	2.98	—	134,681	593	0.04	—
P13-S2	76,930	79,469	300	3.3	—	77,438	585	0.66	—
P13-S3	96,162	99,306	304	3.27	—	96,998	520	0.87	—
P14-S1	126,490	130,247	304	2.97	—	126,631	593	0.11	—
P14-S2	231,898	236,258	311	1.88	—	232,084	607	0.08	—
P14-S3	210,816	216,993	338	2.93	—	211,386	680	0.27	—
P15-S1	186,502	193,141	346	3.56	—	187,353	714	0.45	—
P15-S2	157,810	164,154	379	4.02	—	158,299	741	0.31	—
P15-S3	143,463	148,714	384	3.66	—	143,765	612	0.21	—
P16-S1	114,984	117,376	402	2.08	—	115,975	616	0.86	—
P16-S2	197,115	205,788	394	4.4	—	197,253	768	0.07	—
P16-S3	164,262	170,175	387	3.6	—	168,014	717	2.28	—
P17-S1	14,268	14,772	415	3.53	—	14,321	607	0.37	—
P17-S2	153,992	161,199	420	4.68	—	155,578	748	1.03	—
P17-S3	139,992	145,088	477	3.64	—	141,026	925	0.73	—
P18-S1	31,756	33,007	471	3.94	—	32,351	950	1.87	—
P18-S2	57,698	60,277	483	4.47	—	58,206	837	0.88	—
P18-S3	72,122	74,055	494	2.68	—	72,915	919	1.09	—
P19-S1	119,874	124,909	468	4.2	—	120,979	906	0.92	—
P19-S2	175,140	182,548	482	4.23	—	179,448	761	2.46	—
P19-S3	145,950	151,423	502	3.75	—	146,111	767	0.11	—
P20-S1	105,587	109,008	503	3.24	—	107,456	802	1.77	—
P20-S2	178,545	185,508	500	3.9	—	182,866	776	2.42	—
P20-S3	151,403	156,838	515	3.59	—	155,832	980	2.93	—
		E-GAP (Max) = 4.47% E-GAP(Min) = 0% E-GAP(Mean) = 2.95% S-GAP(Max) = 7.09% S-GAP(Min) = 2.12% S-GAP(Mean) = 4.47%				E-GAP(Max) = 2.93% E-GAP(Min) = 0% E-GAP(Mean) = 0.68% S-GAP(Max) = 3.93% S-GAP(Min) = 0.86% S-GAP(Mean) = 2.4%			

Table 5
Comparison between the solution algorithms for medium- and large-sized problem instances.

Tests			GA			IGA		
	$\|V\|$	$\|P\|$	Objective value	CPU time (S)	AII (%)	Objective value	CPU time (S)	AII (%)
P21-S1	150	10	34,707	1342	11.37	30,839	1724	11.29
P21-S2			28,342	1208	14.13	24,957	1772	12.11
P21-S3			35,323	1309	11.70	32,175	1742	10.43
P22-S1	200	12	51,964	1214	14.60	48,140	1535	14.98
P22-S2			79,938	1243	14.95	70,213	1754	12.4
P22-S3			68,501	1184	12.47	63,532	1435	12.47
P23-S1	250	12	93,396	1301	8.87	83,996	1882	8.87
P23-S2			123,587	1271	10.60	109,415	1721	11.79
P23-S3			148,246	1449	9.49	132,748	2113	9.41
P24-S1	300	14	121,094	1545	9.77	108,593	2380	11.96
P24-S2			89,272	1704	11.97	80,934	2518	8.71
P24-S3			98,433	1292	8.79	86,431	1746	9.89
P25-S1	400	14	53,553	1663	10.61	48,438	2071	12.72
P25-S2			62,064	1764	11.80	56,772	2584	8.54
P25-S3			77,737	1448	13.03	69,618	1730	10.86
P26-S1	450	14	106,355	2084	7.70	97,041	2990	11.56
P26-S2			115,859	2012	9.19	101,287	2991	8.76
P26-S3			102,280	1626	8.24	91,418	2267	11.34
P27-S1	500	16	59,880	2208	11.01	53,795	3266	9.18
P27-S2			152,958	2165	7.48	133,277	3099	10.29
P27-S3			229,469	1930	10.47	207,020	2971	9.52
P28-S1	550	18	864,332	1990	8.52	770,894	3136	8.52
P28-S2			729,018	2044	9.39	629,672	2516	7.68
P28-S3			604,436	1898	9.70	539,597	2645	10.54
P29-S1	600	20	647,801	2156	8.30	552,276	2621	8.35
P29-S2			862,803	2149	7.02	777,024	2594	8.59
P29-S3			707,950	2180	7.51	611,065	2942	9.18
P30-S1	650	20	591,029	2297	6.16	529,278	2863	5.48
P30-S2			788,453	2500	6.27	668,998	3754	5.57
P30-S3			970,268	1982	5.92	870,096	2352	8.15
P31-S1	700	20	1,296,375	2499	6.82	1,158,853	3358	7.44
P31-S2			1,156,451	2558	6.71	1,034,803	3650	7.38
P31-S3			1,265,944	2133	7.46	1,097,442	2659	5.43
P32-S1	750	22	809,855	2593	5.04	681,721	3467	6.94
P32-S2			1,197,700	2507	6.52	1,058,566	3221	7.11
P32-S3			1,087,296	2544	6.69	951,030	3445	6.02
P33-S1	800	22	1,319,604	2667	6.25	1,118,625	3662	5.59
P33-S2			748,538	2404	4.70	656,960	3497	6.27
P33-S3			907,364	2717	6.55	780,727	3921	4.36
P34-S1	900	22	1,328,809	2376	5.78	1,150,786	3156	6.01
P34-S2			2,398,604	2400	5.86	2,077,497	3322	6.72
P34-S3			2,153,248	2375	6.08	1,885,206	3227	5.57
P35-S1	1000	22	2,055,175	2607	4.47	1,772,382	3263	4.22
P35-S2			1,742,781	2356	5.60	1,487,170	3019	5.13
P35-S3			1,564,803	2560	5.70	1,307,062	3072	6.21
P36-S1	1400	24	1,304,759	2411	3.44	1,121,241	3214	4.21
P36-S2			2,208,909	2792	3.74	1,922,034	4029	3.83
P36-S3			1,831,378	2593	3.57	1,525,848	3192	4.89
P37-S1	1500	24	167,533	2499	3.54	146,656	3667	2.57
P37-S2			1,784,236	2632	3.93	1,514,747	3241	2.86
P37-S3			1,611,058	2581	3.50	1,413,112	3059	3.22
P38-S1	1600	24	382,601	2508	2.57	326,440	3544	2.31
P38-S2			687,863	2540	3.68	574,800	3809	3.19
P38-S3			859,155	2673	2.78	750,300	3393	3.03
P39-S1	1800	25	1,516,483	2500	1.91	1,275,674	3436	2.13
P39-S2			2,183,074	2481	2.46	1,897,624	3349	3.19
P39-S3			1,796,617	2853	2.71	1,497,726	4052	1.80
P40-S1	2000	25	1,360,768	2478	1.47	1,163,672	3026	2.02
P40-S2			2,296,271	2469	2.06	1,923,669	3142	2.25
P40-S3			1,921,643	2661	1.96	1,577,469	3887	2.35



a) Residential areas



b) Considered connected graph for Iran's residential areas

Fig. 6. Iran's basic units for the healthcare operation at the first level.

considered to be zero in problem instances. To make sure that the developed algorithms have good performance for our optimization problem, the results of the algorithms are compared with the enumeration-based approach (Z_E^*) for problem instances *P1-P20* in which CPLEX solver achieves the optimal solution of the spanning tree-based MIP model and with the optimal solution from Shirabe's model (Z_{SH}^*) for problem instances *P1-P8* where the Shirabe's model is solvable in Table 4. The algorithms' gap in compared with the enumeration-based approach and Shirabe's model are calculated as E-gap $(\%) = \frac{\text{Algorithm's objective} - Z_E^*}{Z_E^*}$ and S-gap $(\%) = \frac{\text{Algorithm's objective} - Z_{SH}^*}{Z_{SH}^*}$, respectively. It should be mentioned that the enumeration-based approach can find the best possible solution that can be achieved by using the presented algorithms through enumerating all shortest-path trees related to a geographical area graph.

By focusing on the results reported in Table 4, we can see the performance of the IGA in compared with the GA is superior meaningfully in terms of the objective value and the applied approaches for the improvement of the GA are efficient. The average, minimum, and maximum values of E-gap and S-gap related to the GA and IGA are also reported in Table 4.

It should be noted that in GA algorithms, we take care of duplicating a value corresponding to the eliminated edges in solution arrays through their search mechanism, and in such a situation, we change duplicated values to the nearest possible value. On the other hand, at each iteration of the proposed algorithms, the individuals that are not feasible in accordance with *condition* (2) and in districting decisions, the distance between two nodes in a district is more than L^{max} are eliminated from the related population. However, we maintain the population number constant in the next iteration based on the best feasible individuals after the search mechanism. In Table 5, the algorithms are examined for medium- and large-sized instances, and through solving the problem, the average percentage of infeasible individuals (AII) form a population is also reported. It should be noted that in Tables 4 and 5, each problem instance is solved with the presented algorithms four times, and the reported objective function and CPU time are related to the best obtained solution. It is worth noting that in reported CPU times, the time of preprocessing stage regarding to the construction of all shortest-path trees and calculation of their length is also considered. For the largest problem instance, the time of preprocessing stage is about 12 min.

As illustrated in Tables 4 and 5, although the CPU time of the IGA is more than the GA, it has meaningful superiority in comparison with the GA in terms of the objective value.

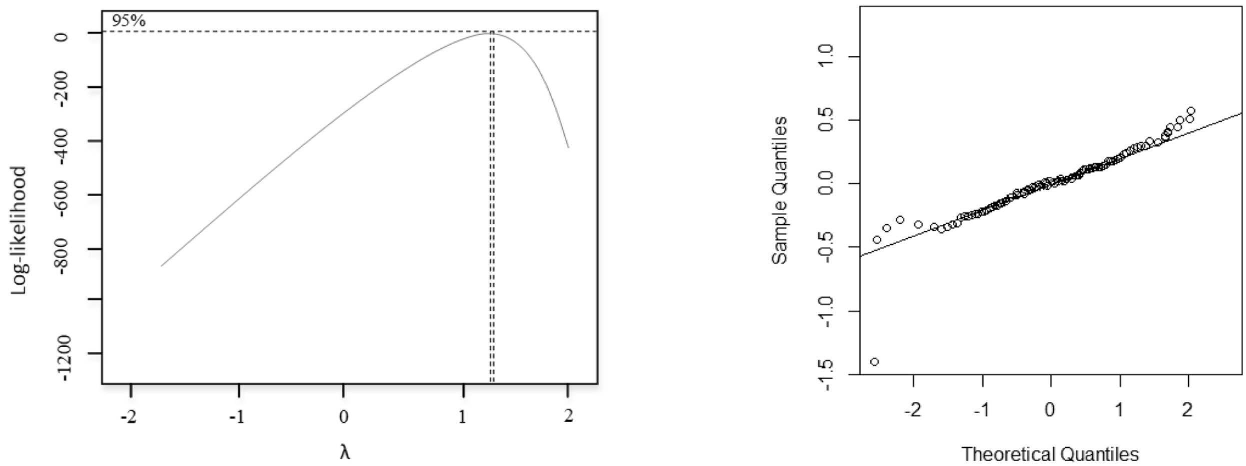
More complementary results, a figure that compares the E-gaps from solving problem instances *P1-P20* by the GA and IGA algorithms and the average CPU time and objective value of instances in four runs of algorithms, are presented in the Supplementary Materials. Finally, based on the presented computational results, we can conclude that the IGA has meaningful superiority in compared with the GA, and the algorithm is suitable for solving the districting problem.

6.2. Case study

Iran with a population of 80 million is 18th-most-populous country in the world with a land area of 1648195 km^2 that is located in the Middle East. The healthcare and medical sector's market value in Iran was almost US $24 billion in 2002 and it is expected to rise to US $96 billion in 2021. Iran healthcare system is an integrated health network that presents three levels of health services as *Level I* that provides specialty and subspecialty medical/paramedical services, *Level II* that includes clinics and hospitals. Doctors and nurses in these units provide health services by using a set of diagnostic facilities such as laboratories, dentistry, radiology, and so on, and *Level III* that consists of health centers such as emergency centers, treatment centers, health homes, and so on that have the most contact with the patient community.

The current infrastructure of Iran's governmental healthcare system is created based on main commitments of the MOHME in providing health services as follows:

1. All residential areas should have access to the third level of health services in <20 km distance,
2. The second and third level of health services should be provided in residential areas with a population size greater than 6000,
3. All service levels should be provided in residential areas with a population size greater than 70,000 and at least one medical sciences university exists in these areas.



a) Sensitivity of the log-likelihood function to the Box-Cox transform parameter λ **b)** Q-Q plot for the Box-Cox transformed model

Fig. 7. Results of Hedonic model for the demand estimation.

In this paper, we focus on partitioning Iran's area for providing the first level of health services. Districting decisions in the first level should be determined by the MOHME. However, the obtained districts from solving the healthcare districting at the first level should be partitioned in the second and third level districting by the major medical science university at the corresponding district. Our case study in this paper is based on the provided information by MOHME related to the provision of governmental health services. In Iran, organizations related to the governmental and private sectors provide health services, and the MOHME as a governmental sector has the most share in providing health services.

Demand and capacity estimation. Case study I is related to defining districts in Iran for the healthcare system operation in the first service level. In this level, in addition to the health system's regulation and management, specialty and subspecialty medical/paramedical services should be provided. In *case study I*, 386 geographically dispersed residential areas are considered for Iran, and it is desired to partition these areas into 10 districts. Further, only 73 areas provide the health services related the first level. Fig. 6 shows the residential areas in Iran and the existing links between them. In Fig. 6(a), the residential areas that provide the corresponding health services are illustrated by red points.

A Hedonic model with Box-Cox transform structure is used for the healthcare demand estimation related to the first level of services. We have gathered the related data from 107 basic units corresponding to the first level of healthcare operation from the MOHME. To capture the seasonality of the demand pattern a weighted average of demand over one year is considered as the real value of the demand. The cumulative demand for the health services at the first level can be estimated by the healthcare systems based on the *referral to hospitals* for these services from doctors through a basic unit. The main information related to gathered data is reported in Table 6.

We have used R software for the estimation of hedonic regression model's parameters. The sensitivity of the log-likelihood function to Box-Cox transform parameter λ is illustrated in Fig. 7(a), and λ is set to be 1.43. Further, Q-Q plot for the Box-Cox transformed model, Fig. 7(b), shows its conformity by normal distribution function.

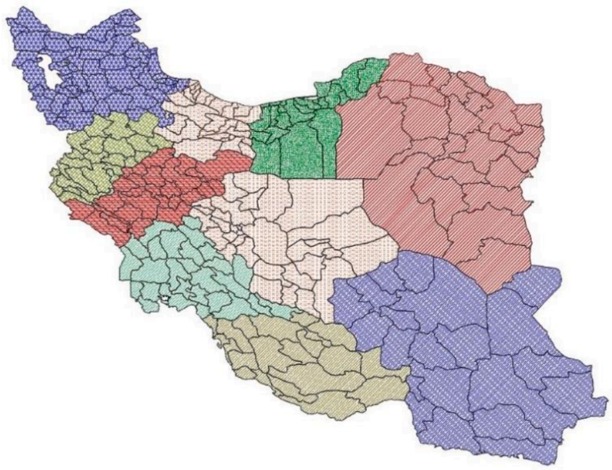
The estimation values of hedonic regression model's parameters are reported in Table 7. The variables with meaningful impacts on the demand value are reported in Table 7. The obtained approximated demand values related to *Case study I* are reported in the Supplementary Materials.

Table 6
The information regarding available samples (107 basic units).

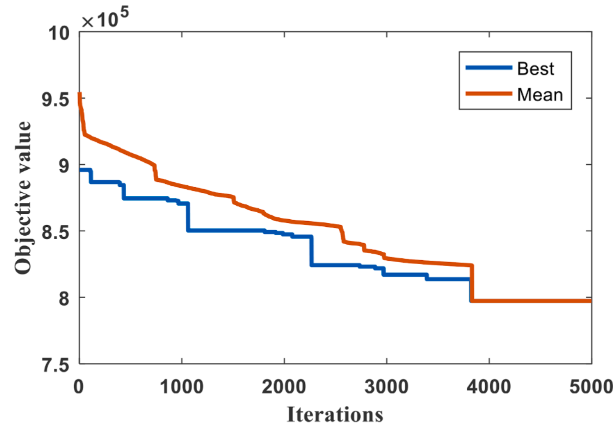
Parameters and variables	Explanations	Mean	Std. deviation
D	Demand value for the first level of health services	19,041	6.74
X_1	The population aging as the number of population more than 65 years old (%)	7.21	2.14
X_2	The percentage of married women between 18 and 45 years old (%)	43.81	11.19
X_3	The unemployment rate (%)	21.46	7.12
X_4	The average of income (Million Rials (Iran's currency))	43.71	22.345
X_5	The literacy level (%)	94.12	4.63
X_6	The urban development (%)	63.19	17.43
X_7	The men' percentage (%)	50.11	1.05
X_8	The population density (population number/KM^2)	332.17	219.84
X_9	The environmental factors (%)	27.14	9.66

Table 7
The outputs of Hedonic model.

	X_1	$(X_1)^2$	X_2	$(X_2)^2$	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Coefficient	0.319	0.0127	0.511	0.319	0.0017	0.00228	0.00084	0.0128	0.0474	0.0223	0.00124
t-stat	2.33**	1.84**	3.28*	1.74**	6.26*	1.78**	9.37**	3.99*	10.34**	10.43**	2.12**
R^2 = 86.31%,											
* $P < 0.05$ ** $P < 0.01$											



a) Solution of districting problem



b) Convergence of IGA for the case study

Fig. 8. Solution of case study *I* with $|P| = 10$.

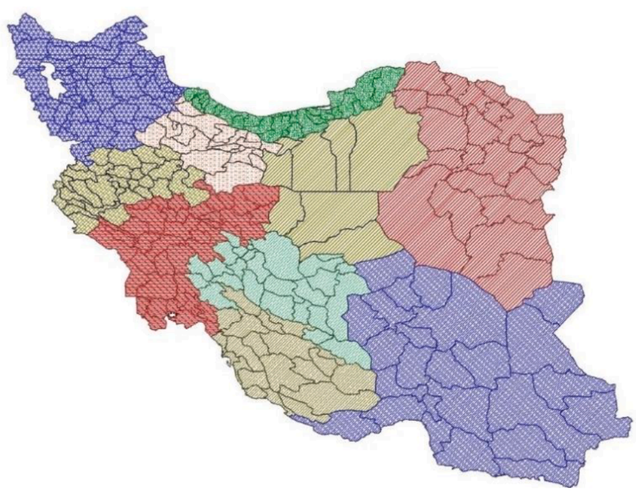


Fig. 9. Iran's current state regarding the health districting in the first level.

The healthcare system in residential areas provides a large number of various health services in the first level and to make the healthcare districting problem possible as a strategic problem, we have assumed that the amount of the existing capacity for various services in *level I* is symmetric to the corresponding demand for them. In other words, the MOHME is committed to providing the infrastructure and capacity for health services based on the demand that can be approximated according to the historical data. For example, if the demand for heart surgery is three times as high as brain surgery, the existing capacity, including surgeons and operating rooms, for heart surgery is also three times higher than the capacity for brain surgery. Further, the healthcare services at the first level contain specialty and subspecialty medical services that should be provided in hospitals. The MOHME approximates the capacity for providing health services at the first level based on the available facilities at hospitals of basic units such as Hospitals' beds, ICU beds, CCU beds, etc (http://ird.behdasht.gov.ir/). The data related to the capacity of areas in *Case Study I* is also reported in the Supplementary Materials. Therefore, it is possible to deal with the healthcare districting problem at the first level of healthcare services by considering the cumulative capacity and demand in residential areas. Based on the approximated demand, in various residential areas, the percentage of the population that requires services in *level I* is within [9, 24] %.

Districting decisions. In case study *I*, $|V|$=386, $|P|$=10, and L^{max} is assumed to be 800 KM. Information related to the traveling distance between basic units are obtained from the Iran's Road Maintenance & Transportation organization. Case study *I* is solved for $\lambda = 0$ using the GA and IGA, and the IGA has a better performance for the case study. The convergence of the IGA is illustrated in Fig. 8(b) in which the red line shows the average fitness value of the population and the blue line illustrates the best fitness value in the population. The obtained solution for case study *I* from the IGA is also shown in Fig. 8(a).

In Fig. 9, the current state of Iran's districts for provision of health services in the first level is presented. By analyzing our solution in compared with the current districts, it is observed that our solution reduces the balance criterion, the problem's objective function, from 1,175,419 to 797,274 that is equal to 32% improvement. Further, in our solution the capacity of only one district is less than the estimated demand for the corresponding health services, but in three districts of the current state, the capacity level is less than the demand.

In the case study, the defined balance criterion as the problem's objective function reduces the number of districts in which the capacity level for the service provision is less than the demand. By reducing these districts, the total number of patients that should travel to other districts for accessing to services decreases.

Hierarchical healthcare districting approach. The structure of Iran's healthcare system that includes three levels of healthcare services with various decision makers for districting at different levels leads us to propose a hierarchical districting approach. There exists a hierarchical structure among the three levels of healthcare services and based on our collaborations with the MOHME, obtained districts from solving the districting problem at the first level should be partitioned in the second/third level healthcare districting. It should be mentioned that depending on the context, there might be different concerns and criteria for the

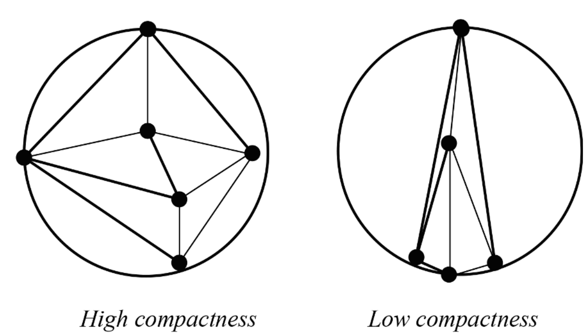


Fig. 10. Illustrative examples of high and low compactness.

Table 8
The sensitivity of the objective function, the compactness measure, and AWD to L^{max} and $|P|$ when $\lambda = 0$.

L^{max}		$\vert P\vert = 10$	$\vert P\vert = 20$	$\vert P\vert = 30$	$\vert P\vert = 50$	$\vert P\vert = 60$	$\vert P\vert = 70$	$\vert P\vert = 80$	$\vert P\vert = 90$	$\vert P\vert = 100$
800	Objective value	797,274	394,040	219,926	144,699	144,241	143,284	126,449	116,070	92,340
	Min value of *CM* over all districts	0.18	0.11	0.14	0.17	0.20	0.11	0.17	0.19	0.11
	Max value of *CM* over all districts	0.42	0.37	0.36	0.53	0.48	0.34	0.40	0.51	0.57
	Mean value of *CM*	0.34	0.16	0.25	0.22	0.24	0.25	0.27	0.24	0.24
	Min value of *AWD* over all districts	50	37	31	30	30	27	25	18	14
	Max value of *AWD* over all districts	103	85	65	42	67	58	52	41	31
	Mean value of *AWD*	63	48.10	39.99	39.60	39	34.56	32.75	21.96	19.46
650	Objective value	-*	531,581	377,184	201,105	152,162	151,796	130,303	118,831	104,904
	Min value of *CM* over all districts	–	0.16	0.18	0.18	0.18	0.12	0.17	0.15	0.18
	Max value of *CM* over all districts		0.29	0.35	0.29	0.57	0.42	0.48	0.52	0.53
	Mean value of *CM*	–	0.48	0.19	0.44	0.33	0.20	0.31	0.21	0.21
	Min value of *AWD* over all districts	–	32	27	29	28	25	24	17	14
	Max value of *AWD* over all districts	–	75	62	40	60	52	50	39	30
	Mean value of *AWD*	–	41.37	35.99	35.24	35.10	31.45	27.84	19.32	18.10
450	Objective value	–	649,667	393,802	318,472	202,703	199,292	165,782	121,878	105,253
	Min value of *CM* over all districts		0.17	0.14	0.26	0.10	0.24	0.25	0.19	0.14
	Max value of *CM* over all districts	–	0.38	0.28	0.57	0.21	0.52	0.50	0.38	0.32
	Mean value of *CM*	–	0.25	0.21	0.40	0.15	0.36	0.39	0.27	0.23
	Min value of *AWD* over all districts	–	29	25	26	27	22	23	16	13
	Max value of *AWD* over all districts	–	66	59	38	55	46	47	36	27
	Mean value of *AWD*	–	36.82	33.83	32.07	29.84	27.36	25.33	16.62	16.83
250	Objective value	–	841,522	681,833	428,159	291,323	280,394	213,471	187,207	174,330
	Min value of *CM* over all districts	–	0.21	0.20	0.26	0.17	0.18	0.25	0.19	0.20
	Max value of *CM* over all districts	–	0.41	0.45	0.58	0.34	0.40	0.50	0.40	0.42
	Mean value of *CM*	–	0.31	0.29	0.42	0.27	0.27	0.35	0.30	0.30
	Min value of *AWD* over all districts	–	26	22	24	24	19	22	16	13
	Max value of *AWD* over all districts	–	61	51	34	47	41	43	31	24
	Mean value of *AWD*	–	34.24	31.80	29.83	28.04	24.63	22.55	14.63	15.82
100	Objective value	–	–	–	–	367,268	345,753	313,721	200,691	193,696
	Min value of *CM* over all districts	–	–	–	–	0.19	0.25	0.18	0.18	0.19
	Max value of *CM* over all districts	–	–	–	–	0.43	0.55	0.41	0.40	0.39
	Mean value of *CM*	–	–	–	–	0.30	0.39	0.28	0.28	0.30
	Min value of *AWD* over all districts	–	–	–	–	23	17	21	14	12
	Max value of *AWD* over all districts	–	–	–	–	42	36	40	30	21
	Mean value of *AWD*	–	–	–	–	24.12	21.92	19.61	12.43	14.56

*In instances corresponding to these rows, IGA couldn't find any feasible solution.

Fig. 11. Solution of case study *I* with $|P| = 30$ and $L^{max} = 800$.

districting problem at various levels and in this study, we have focused on the first level of services. The proposed approach in this sub-section can be used for the districting at the second/third level of services after the consideration of their criteria and concerns in the mathematical model.

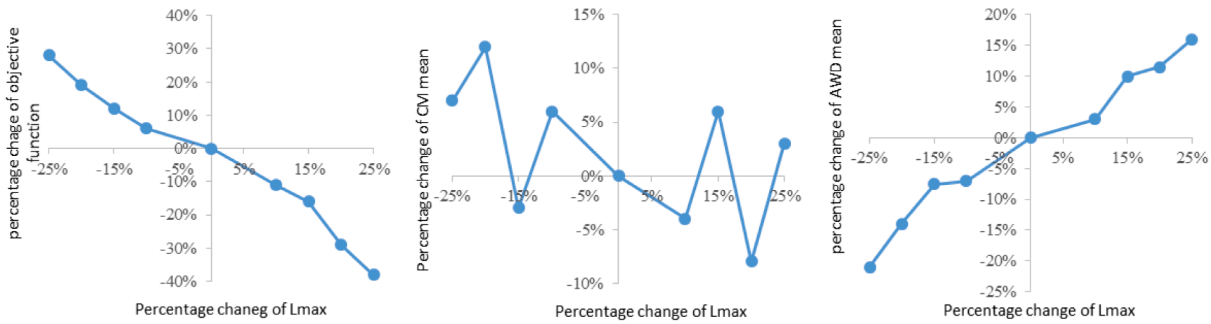
For this purpose, we propose a *coarsening* procedure that creates multiple hierarchical levels of graphs. In this procedure, adjacent nodes are merged to create smaller graph in the first/second level districting and then, for the districting at the next level, we map the obtained solution in the upper level by the reduced graph instance back to the original graph. The proposed coarsening procedure is explained in Algorithm C1 in Appendix C. The idea behind the coarsening procedure is to 1- reduce the computational intractability of the optimization problem by decreasing the number of basic units, 2- improve the solution of presented algorithms, and 3- make the districting problem solvable by MIP-based approaches in some cases.

By using the coarsening procedure, the solution of the IGA has been slightly improved in some large-sized instances and *Case Study I*. Further, by the proposed approach, the CPU time of the IGA is reduced, meaningfully. It is worth noting in the case study, we implemented this procedure and the original graph with 1021 nodes is reduced to presented graph in Fig. 6 with 386 nodes.

Although there might be different concerns and criteria for the districting problem at the third level, we have examined a case example to only illustrate the implementation of the hierarchical districting approach in the Supplementary Materials. We consider one of the obtained districts from solving case study, and after mapping back the district to the original graph, we partition it for the healthcare provision at the third level by the proposed mathematical model.

6.3. Analysis of results and discussion

In this section, in addition to the objective function that represents the balance of districts, we consider the compactness of districts and average weighted distance of the population to the healthcare services to analyze the impact of problem's parameters on the quality of districts.



a) Objective function **b) CM mean** **c) AWD mean**

Fig. 12. Sensitivity of main measures corresponding to districting decisions to L^{max}.

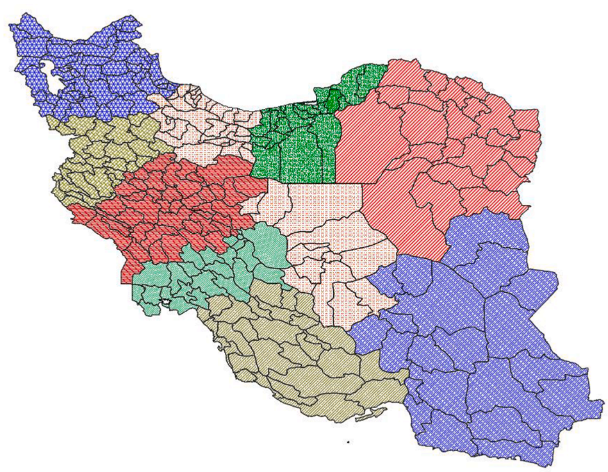


Fig. 13. . Solution of the case study by using the alternative fitness function.

Compactness. we can evaluate the geometrical shape of a district by a compactness measure, and there are several ways to quantify it in the related literature. Although there is not any restriction on the compactness of districts in our optimization problem, we investigate the compactness of obtained districts from solving the problem. For district D_p related to graph G, the compactness means that the convex hull induced by $V_{D(p)}$ shall be nearly round-shaped, undistorted, and has a smooth boundary. Let G be our geographical area containing all basic units, and C_p be the smallest circle enclosing the basic units of D_p, then based on (Datta et al., 2013; and Kalcsics and Ríos-Mercado, 2019), we calculate the compactness of district D_p, CM_p, as $CM_p = \frac{AR(D_p)}{AR(G \cap C_p)}$, where $AR(D_p)$ denotes the area of district D_p, and $AR(G \cap C_p)$ denotes the overlap area of C_p and G. Based on the defined measure, Fig. 10 represents districts with high and low compactness. Further in the Supplementary Materials, the compactness of obtained districts are reported for several instances. The compactness of districts calculated using ArcGIS 10.6 software.

Average weighted distance of the population to the healthcare services. For the first level of healthcare services, the corresponding services are not provided in all basic units. Therefore, the average weighted distance of the population to the healthcare services (AWD) is calculated for each obtained district to analyze the quality of the solution. In district D_p, if we denote the distance of basic unit i with the demand size of α_i to the nearest residential area that provides the healthcare services as θ_i, then the AWD can be calculated as $AWD_p = \frac{\sum_{i \in D_p} \alpha_i \theta_i}{\sum_{i \in D_p} \alpha_i}$.

The districting problem's solution is dependent on the value of parameter L^{max} and the number of desired districts. In the case study, Table 8 presents the sensitivity of the obtained objective function by the IGA, the compactness measure, and AWD to L^{max} and $|P|$ when $\lambda = 0$. We have also presented the obtained solution from solving the case study by the IGA for $|P| = 30$ and $L^{max} = 800$ in Fig. 11, and in the related districting solution, the capacity level is more than the demand in all districts. We have also investigated the impact of L^{max} on the objective value and mean values of CM and AWD by using some problem instances, *P15*, *P18*, *P20*, *P24*, *P27*, *P30*, *P33*, *P36*, and *P40* from data set *S1*, and results related to the average impact of L^{max} on mentioned measures are shown in Fig. 12.

From the presented results in Table 8 and Fig. 12, we can conclude:

- The value of parameter L^{max} has a main impact on the feasible region of the districting problem. Decreasing the value of this parameter reduces the feasible region of the problem and increases its objective function value.
- By increasing the number of districts, it is possible to reduce the objective function and obtain the districts with more suitable shapes.

Table 9
The districting decisions' characteristics related to the primary objective, alternative objective, and current state.

	IGA districting solution		Current districting
	Alternative objective	Primary objective	
Primary objective value	861,056	797274*	1,175,419
Alternative objective value	0.097*	0.114	0.217
Min value of *CM* over all districts	0.21*	0.18	0.16
Max value of *CM* over all districts	0.48*	0.42	0.45
Mean value of *CM*	0.35*	0.34	0.29
Min value of *AWD* over all districts	54	50	38*
Max value of *AWD* over all districts	105	103*	144
Mean value of *AWD*	65.2	63*	68.5
Number of districts with shortage capacity	1*	1*	3

*best value

Moreover, for small desired number of districts, it is not possible to find any feasible solution for the problem with small value of L^{max}.

- The decrease of parameter L^{max} and increase of the desired number of districts, $|P|$, result in districts with smaller areas and improve the *AWD* value of districts.
- Based on the presented results, the values of L^{max} and $|P|$ do not have any significant impact on the considered compactness measure.

In this sub-section, it is highlighted that both the decrease of L^{max} and increase of $|P|$ improve the accessibility of the patients to the healthcare services and the increase of $|P|$ and L^{max} improve the balance criterion. However, the values of these predetermined parameters are related to the management level of a healthcare system, and increasing the number of districts may increase the difficulty of the healthcare management.

Alternative objective function. In order to make the capacity-demand balance, we define an alternative objective function for the IGA based on the capacity usage ratios at the districts as follows:

$$min: \frac{1}{|P|}\sum_{p\in P}\left(\frac{\sum_{i\in V}\alpha_i x_{ip}}{\sum_{i\in V}\delta_i x_{ip}} - 1\right)^2 \tag{20}$$

Objective function (20) makes the capacity ratios at the districts close to one and increase the balance of districts. However, this function is non-linear and non-convex and using it as mathematical models' objective makes them computationally intractable. The obtained solution for the case study by the IGA with fitness function (20) is shown in Fig. 13.

By comparing Figs. 13 and 8(a), the districting solutions have some differences, especially with the districts in the western part of the country. Further, the results of solving Case Study with these two objectives and the current state of districting decisions are compared in Table 9 in terms of the average, minimum, and maximum values of *CM* and *AWD* over all obtained districts.

Based on the results reported in Table 9, the values of *CM* measure for the alternative objective function are relatively better than the primary one, although the *AWD* measure's values have superiority by applying the primary objective function. By using both of these objectives, the number of districts with the shortage capacity is the same. Further, we can conclude that applying both of objectives improves the quality of districting solutions in compared with the current districting state, meaningfully. In Table 10, the results from solving the case study by the IGA with this new fitness function are summarized for various values of L^{max} and $|P|$.

By comparing the results from solving the instances with primary and alternative objectives, reported in Tables 9 and 10, we can conclude that the differences between the values of the *CM* and *AWD* measures of districting solutions are not significant. Further, in various values of L^{max} and $|P|$, the existing trend of the alternative objective function is relatively the same as the primary objective function.

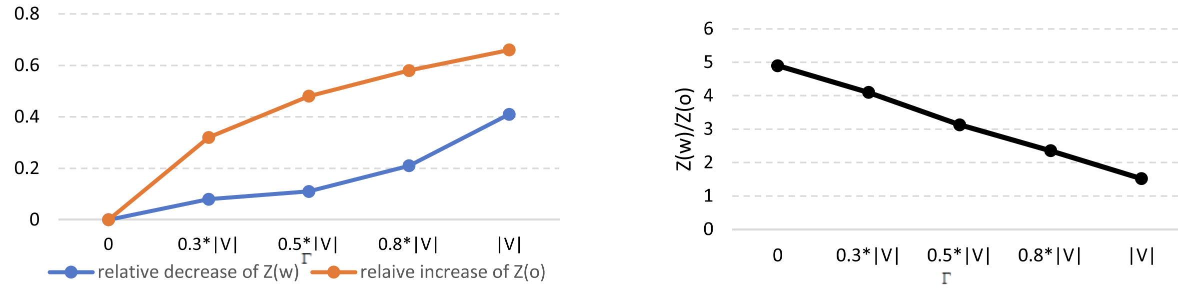
6.4. Analysis of the robustness consideration

As mentioned before, the approximation of demand for health services in residential areas is a challenging task and it is difficult to have a precise estimation for it. On the other hand, the demand value approximation is crucial to properly plan for a healthcare system (Harper et al., 2005; Feldstein, 1966). In this study, the interval uncertainty is employed to deal with the demand uncertainty, and in each residential

Table 10
The sensitivity the compactness measure and *AWD* to L^{max} and $|P|$ by the alternative objective function.

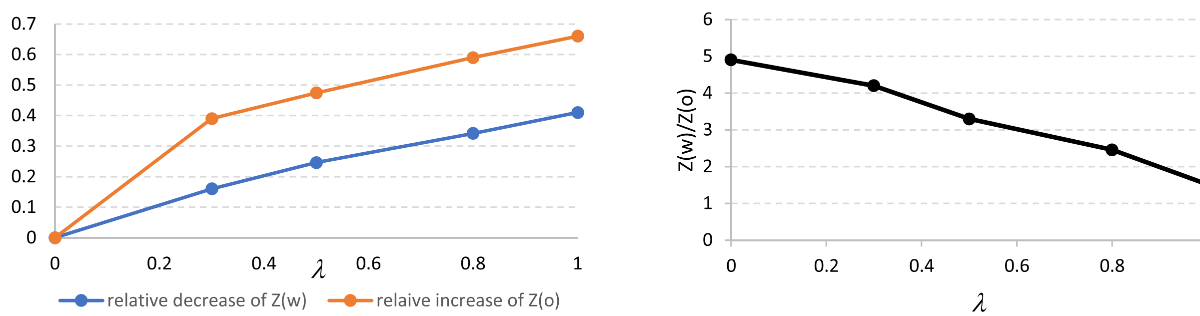
L^{max}		$\|P\| = 10$	$\|P\| = 20$	$\|P\| = 30$	$\|P\| = 50$	$\|P\| = 60$	$\|P\| = 70$	$\|P\| = 80$	$\|P\| = 90$	$\|P\| = 100$
800	Alternative objective value	0.097	0.09	0.082	0.077	0.075	0.068	0.049	0.044	0.041
	The value of primary objective function	861,056	413,742	230,923	156,275	165,878	150,449	139,094	123,035	100,651
	Mean value of *CM*	0.35	0.17	0.29	0.19	0.25	0.21	0.30	0.23	0.28
	Mean value of *AWD*	65.2	47.37	39.73	37.6	33	27.68	26	25.7	20.1
650	Alternative objective value	-*	0.087	0.081	0.064	0.06	0.05	0.036	0.037	0.027
	The value of primary objective function	–	584,740	411,131	219,205	171,944	160,904	140,728	131,903	117,493
	Mean value of *CM*	–	0.24	0.27	0.27	0.31	0.29	0.28	0.25	0.26
	Mean value of *AWD*	–	43.50	40.34	35.59	32.99	28.61	23.10	16.22	20.63
450	Alternative objective value	–	0.082	0.078	0.067	0.054	0.044	0.038	0.034	0.024
	The value of primary objective function	–	740,621	452,873	366,243	214,866	217,229	184,019	138,941	116,831
	Mean value of *CM*	–	0.29	0.26	0.22	0.32	0.24	0.23	0.17	0.24
	Mean value of *AWD*	–	37.18	36.87	28.22	31.03	28.45	24.82	19.2	18.47
250	Alternative objective value	–	0.052	0.042	0.051	0.049	0.043	0.037	0.031	0.021
	The value of primary objective function	–	942,505	736,380	458,131	320,456	302,826	230,549	211,544	196,993
	Mean value of *CM*	–	0.25	0.30	0.28	0.29	0.29	0.28	0.27	0.29
	Mean value of *AWD*	–	35.03	35.66	27.02	25.63	21.67	20.29	16.48	15.61
100	Alternative objective value	–	–	–	–	0.046	0.042	0.036	0.028	0.021
	The value of primary objective function	–	–	–	–	385,632	366,499	338,819	230,795	220,814
	Mean value of *CM*	–	–	–	–	0.28	0.25	0.23	0.21	0.23
	Mean value of *AWD*	–	–	–	–	23.70	20.19	19.08	15.79	14.30

*In instances corresponding to these rows, IGA couldn't find any feasible solution.



a) the relative increase and decrease of z^O and z^W in various values of Γ b) The changes of $\left(z^W/z^O\right)$ in various values of Γ

Fig. 14. The performance of the BSRO.



a) the relative increase and decrease of z^O and z^W in various values of λ b) The changes of $\left(z^W/z^O\right)$ in various values of λ

Fig. 15. The performance of the WWRO.

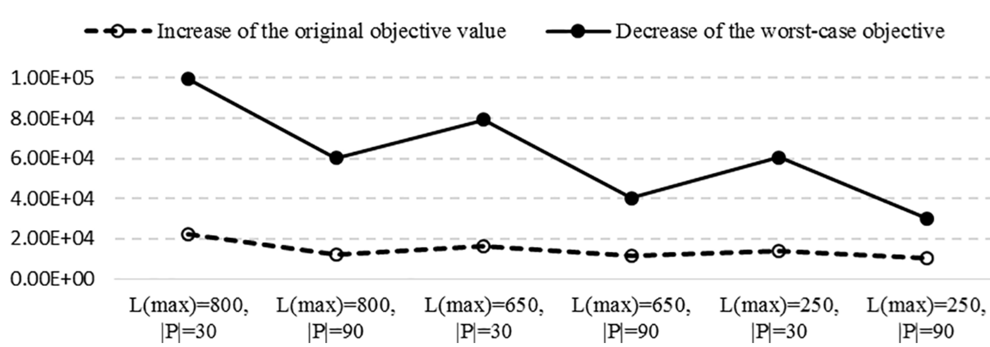


Fig. 16. The comparison between the cost and benefit of robustness consideration in the case study by WWRO method.

area, the nominal value of the demand for health services is approximated as a percentage of the population size and its variation in relative to its nominal value is also considered. In this section, the impact of robustness consideration is investigated.

Firstly, we examine BSRO and the WWRO for some problem instances. To do so, in the MIP model, the objective function corresponding to optimal solution $\boldsymbol{x}^*$ in BSRO and WWRO approaches are denoted by z^*. In addition, z^O denote the original objective value of $\boldsymbol{x}^*$ in BSRO and WWRO approaches, which are calculated as $\max_{p,p' \in P}\left\{\sum_{i\in V}\left(\delta_i-\overline{\alpha}_i\right)x^*_{ip}-\sum_{i\in V}\left(\delta_i-\overline{\alpha}_i\right)x^*_{ip'}\right\}$. The increase of the original objective function can be interpreted as the cost of robustness. The worst-case value of objective related to optimal solution $\boldsymbol{x}^*$, $\max_{p,p' \in P}\left\{\sum_{i\in V}\left(\delta_i-\overline{\alpha}_i\right)x^*_{ip}-\sum_{i\in V}\left(\delta_i-\overline{\alpha}_i\right)x^*_{ip'}+\left(\sum_{i\in V}\widehat{\alpha}_i\left(x^*_{ip}+x^*_{ip}\right)\right)\right\}$, is denoted by z^W.

To have a robust solution, we increase the value of control parameter Γ, $0\leqslant\Gamma\leqslant|V|$, in BSRO and λ, $0\leqslant\lambda\leqslant1$, in the WWRO approach. The cost of robustness is the increase of original objective value by increasing these parameters, and its benefit is the reduction of the worst-case objective. If the benefit of the robustness consideration be meaningful in compared of its cost, we can conclude that the robustness consideration is appropriate for our problem. By solving the districting problem using the enumeration-based approach, for the several problem instances with different values of λ and Γ, the average changes of z^O and z^W are shown in Figs. 14 and 15 for BSRO and WWRO approaches, respectively.

Our experimental results in Figs. 14 and 15 illustrate that setting

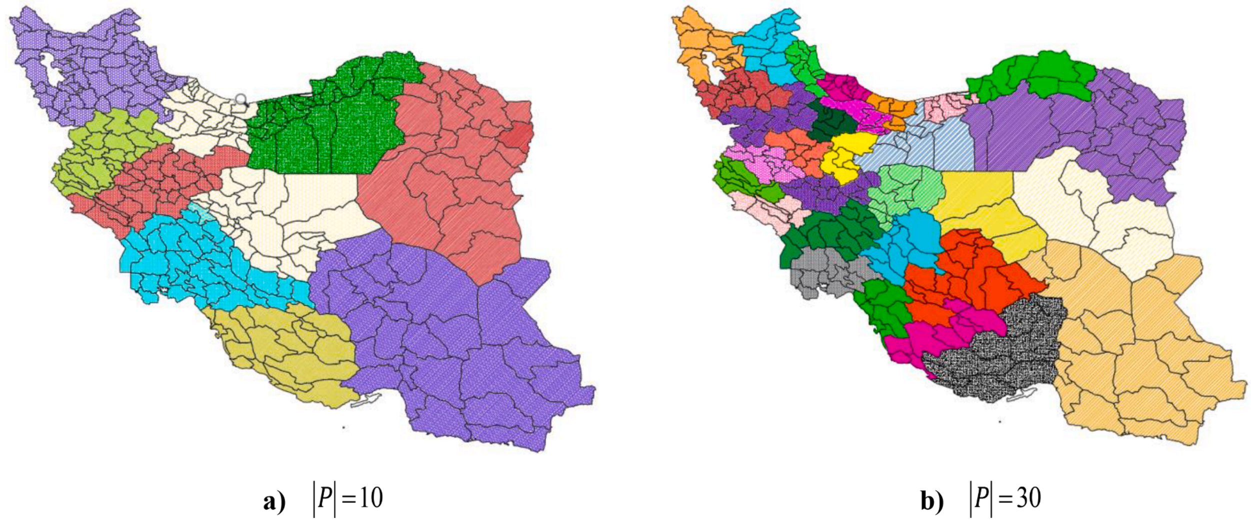


Fig. 17. Robust solution of the case study when $\lambda = 1$.

higher values for control parameters Γ and λ leads to higher values for the original objective and lower values for the worst-case objective in problem instances. However, the meaningful decrease of the worst-case objective, the benefit of the robustness consideration, in comparison with the increase of the original objective, the cost of the robustness consideration, highlights the importance of the robust solution. For the highest protection against the uncertainty, the values of original objective, worst-case objective, and the problem's objective are the same in both of approaches and these approaches allow for considering different risk attitudes of a decision maker by carefully adaption of control parameters. In addition, the trend of values of z^*, z^O, and z^W in compared with the control parameters' values are relatively the same for both of approaches. In terms of CPU time, the WWRO approach has significant superiority in compared with BSRO approach and we can use the defined robustness measure based on the WWRO approach as the fitness function of meta-heuristics. Because of these issues, we have used the WWRO approach in the rest of paper.

To emphasize on the applicability of the proposed WWRO approach for the robustness consideration, we have solved the case study by the IGA with various values for L^{max} and $|P|$ when $\lambda = 1$, and the interval uncertainly for each basic unit is uniformly generated between [10, 30] % of $\overline{\alpha}_i$. The increase of the original objective value is compared with the decrease of the worst-case objective in Fig. 16.

Fig. 16 highlights the decrease amount of the worst-case objective value in comparison with the increase of the original objective value is meaningfully larger that shows the applicability of our approach. Further, in the case study, the changes of the original and worst-case objective value in obtaining robust solutions are dependent on the problem's setting, such as values of parameter $|P|$ and L^{max}.

By robustness consideration in the optimization problem, the optimal decisions related to partitioning the basic units change, meaningfully. In the case study, by $L^{max} = 800$, we present the robust districts when $\lambda = 1$ in Fig. 17(a) and (b) for $|P| = 10$ and $|P| = 30$, respectively.

By comparing Fig. 17(a) with Fig. 8(a) and Fig. 17(b) with Fig. 11, we can see the number of residential areas in some districts decreases to reduce their corresponding uncertainty of the cumulative demand and in some other districts, the number of assigned residential areas is increased, which increases the overall uncertainty. In other words, the robust solution improves the balance between the existing uncertainty in the cumulative demand of various districts.

7. Conclusion

This study was a new attempt to propose an MIP model for the healthcare districting problem considering balance and contiguity on districts based on the graph-based nature of this problem. The proposed optimization problem is applicable to partition residential areas for a healthcare system operation. The proposed MIP model is based on a predetermined shortest-path tree related to the connected graph of a geographical area. To find a proper solution, the enumeration-based approach is developed in which we solve the MIP model for all corresponding shortest-path trees. By computational results, we have shown although the enumeration-based approach has about 2.1% average optimality gap for generated problem instances, the model has a better performance compared to the presented model based on (Shirabe, 2005, 2009) in terms of computational tractability, and our model is solvable for small- and medium-sized instances by CPLEX commercial solver.

This study developed two approaches, BSRO and WWRO, to hedge against the interval uncertainty related to the size of the population in need of medical services and so they enable a decision-maker to obtain robust districting decisions with consideration of different risk attitudes. Our robust solution results in higher values for the original objective while lower values for the worst-case objective. Also, the decrease of the worst-case objective value is more than the increase of the original objective.

Because of the NP-hardness of the districting problem, the GA and its improved version , IGA, are proposed based on the graph-based nature of the problem to solve large-sized problem instances. Computational results are presented for several randomly generated problem instances and a real-world case study in Iran. The computational results demonstrate that IGA has a better performance in terms of the objective function value in comparison with the GA. The optimality gap between the optimal and the approximate solutions, obtained by the IGA, is relatively low in our results.

Our empirical results from case study show that the use of our proposed optimization approach for the healthcare districting problem leads to the balance criterion improvement by 32% in comparison with the existing districting decisions. Further, our solution reduces the number of districts in which the capacity of health services is less than the demand. The results from the case study also show that district-related parameters (here the number of desired districts in a geographical area and the maximum allowable traveling time/distance in districts) have significant effects on the features of districting decisions and the balance among districts. It is shown that the decrease of L^{max} and increase of the desired number of districts decrease the average weighted distance of the population to the healthcare services and the increase of L^{max} improves the balance criterion. Further, the quality of districting solutions from our optimization tools, the importance of robustness consideration, and the hierarchical approach for multi-level districting are investigated in our extensive computational results section.

According to the specialists of Iran's MOHME, the proposed optimization approaches can be used as powerful managerial tools in a healthcare system to provide suitable and more homogeneous services. In a healthcare system related to a country or wide geographical area, the healthcare services at various levels should be provided for the population, and usually, different organizations may be responsible for districting at various levels. Therefore, the idea of *hierarchical* districting helps decision makers partition a geographical area for the provision of health services in different levels.

Our study does not consider a large-scale demand on certain types of healthcare services because of a disease outbreak, such as the spread of coronavirus disease 2019 in the world, and addressing this issue is challenging research in the area of healthcare districting. The multi-level districting and coarsening procedure were not our focus in this study, and future researches can work on this problem and present efficient solution algorithms that coarsen a geographical area graph until its districting problem becomes solvable by exact methods. Lastly, although our optimization tools are motivated by healthcare operations, the proposed tools and insights can be applied to other districting problems in the context of service operations management.

CRediT authorship contribution statement

Sobhan Mostafayi: Software, Visualization, Investigation. **Mohammad Fattahi:** Supervision, Conceptualization, Resources, Project administration. **Esmaeil Keyvanshokooh:** Conceptualization, Validation.

Appendix A. The GA framework

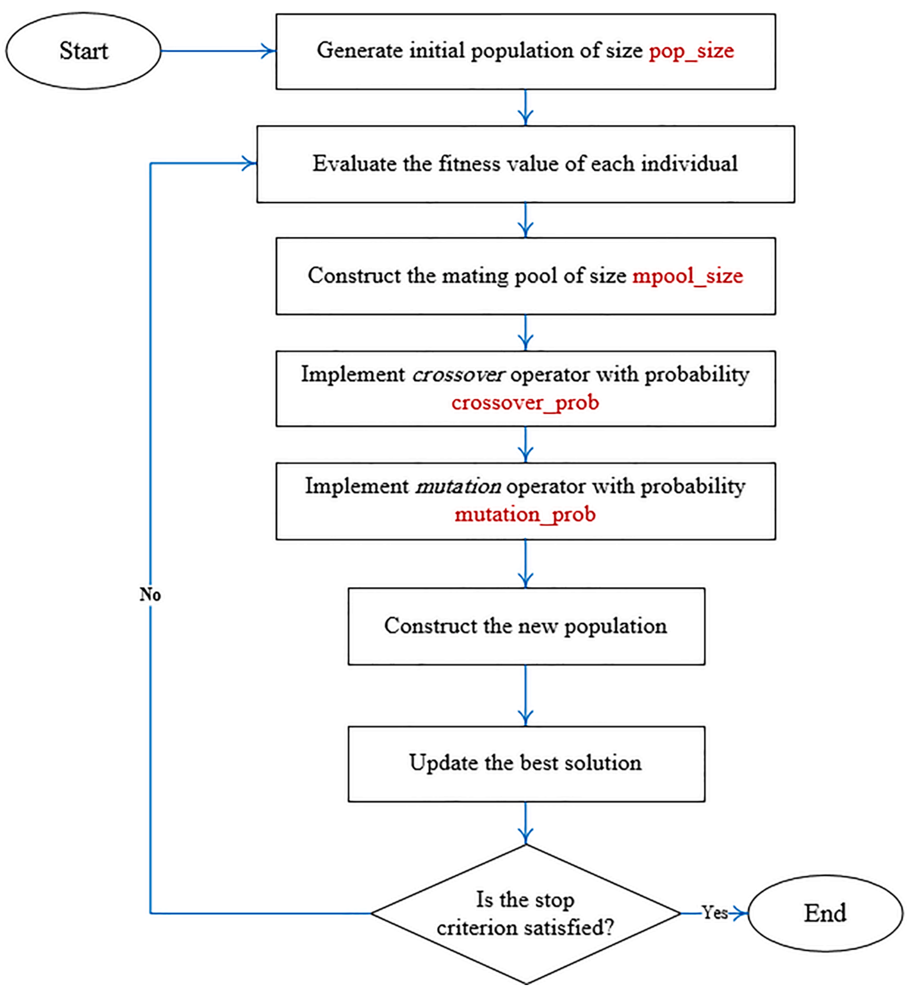


Fig. A1. The flowchart of the GA.

Appendix B:. Generation of problem instances

To present a geographical area for problem instances, an adjacency graph based on (Haugland et al., 2007) is generated. To obtain the adjacency graph $G = (V, E)$ related to each instance, two geographical dimensions x_i and y_i are randomly generated using uniform distribution U [10, 1200] for each vertex $i \in V$. Then, by Algorithm B1, we construct the adjacency graph $G = (V, E)$ from the complete graph $G' = (V, A)$ corresponding to the generated vertices. In the complete graph, the weight of each edge incident to both vertices i and $j \in V$ is based on the Euclidean distance between them,$\tau_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$. In Algorithm B1, we compare pairs of edges in A, and if they intersect, the edge with the most weight is removed from the set (Haugland et al., 2007).

Algorithm B1- *the adjacency graph generation*

Input: the complete graph G' related to generated V and $G' = (V, A)$
  Set $E = A$.
  For $(i,j) \in A$ **do**
    If $(i,j) \in E$ **then**
      For $(u,v) \in E \backslash \{(i,j)\}$ **do**
        If (u,v) and (i,j) intersect *and* $\tau_{uv} \geqslant \tau_{ij}$ **then** remove (u,v) from E;
      End for
  End for
  Set $G = (V, E)$.
Output: G

The obtained graph G from Algorithm B1 is a planar graph. After generating connected graph G, 70% of its vertices are randomly considered as urban nodes, and other ones are assumed to be rural ones. The population of urban and rural nodes are generated using the normal distribution function as $round(N(800, 400) \times 10^3)$ and $round(N(300, 200) \times 10^2)$, respectively. We generate three sets of problem instances as follows:

Problem instances S1: In this set, only 20% of urban nodes present health services, and the services can be provided for U [30, 50]% of the corresponding nodes' population. Further, the nominal percentage of nodes' population that require health services is considered as 8%.

Problem instances S2: In this set, 60% of urban nodes and 10% of rural ones present the health services, and the services can be provided for U [20, 40] % of the corresponding nodes' population. The nominal percentage of nodes' population that require health services is considered as 12%.

Problem instances S3: In this set, all basic units, including urban and rural nodes, present health services, and the services can be provided for U [15, 35]% of the corresponding nodes' population. The nominal percentage of nodes' population that require health services is considered as 25%.

L^{max} is set as the maximum length of shortest-paths in generated graph G multiplied by $(3/|P|)$. In the generated problem instances, for each basic unit, the maximum deviation of patients' number from its nominal value is uniformly generated between U[10, 30]% of the nominal number of patients.

Appendix C:. Coarsening procedure

Algorithm C1- *Coarsening procedure*

Input: connected graph G_0,$G_0 = (V, E)$, related to a geographical area.
Initialize α^c and ℓ^c; Set $V_c = \{i \in V | \alpha_i \leqslant \alpha^c\}$,
While $(|V_c| > 0)$
  Find basic unit i from V_c with the minimum value of α_i,
  Find basic unit j with the nearest distance,τ_{ij}, to basic unit i,
  If $(\tau_{ij} \leqslant \ell^c)$ **then**
    eliminate basic unit i and merge with unit j;
    $\alpha_j \leftarrow \alpha_i + \alpha_j$;
    $\delta_j \leftarrow \delta_i + \delta_j$;
    update E and V,
  End if
  $V_c \leftarrow V_c \backslash \{i\}$,
End while
Output: final graph $G = (V,E)$ after coarsening,

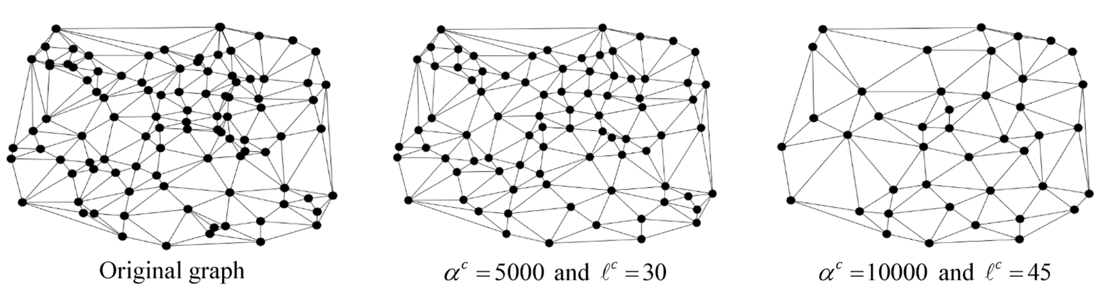


Fig. C1. Coarsening procedure with various values of α^c and ℓ^c.

The reduced graph from the presented coarsening procedure is dependent on values of parameters α^c and ℓ^c. Fig. C1 shows the obtained adjacency graphs from this procedure with various values of input parameters for an illustrative example. For our case study, the value of α^c and ℓ^c are set based on their performance in a large number of empirical experiments using several problem instances as $\alpha^c = 5000$ and $\ell^c = 30$. In the Supplementary Materials, in some large instances, the performance of this procedure is investigated in terms of the CPU time and final objective of the IGA.

Appendix D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cor.2021.105425.

References

Ahuja, R.K., Magnanti, T.L., Orlin, J.B., 1990. Network Flows: Theory. Algorithms Appl.

Akin, J.S., Guilkey, D.K., Hazel, E., 1995. Quality of services and demand for health care in Nigeria: a multinomial probit estimation. Soc. Sci. Med. 40 (11), 1527–1537.

Bacao, F., Lobo, V., Painho, M., 2005. Applying genetic algorithms to zone design. Soft Comput.-A Fusion of Found. Methodol. Appl. 9 (5), 341–348.

Benzarti, E., Sahin, E., Dallery, Y., 2013. Operations management applied to home care services: Analysis of the districting problem. Decis. Support Syst. 55 (2), 587–598.

Bertsimas, D., Sim, M., 2004. The price of robustness. Oper. Res. 52 (1), 35–53.

Blais, M., Lapierre, S.D., Laporte, G., 2003. Solving a home-care districting problem in an urban setting. J. Operat. Res. Soc. 54 (11), 1141–1147.

Bozkaya, B., Erkut, E., Laporte, G., 2003. A tabu search heuristic and adaptive memory procedure for political districting. Eur. J. Oper. Res. 144 (1), 12–26.

Calvete, H.I., Galé, C., Iranzo, J.A., 2016. MEALS: A multiobjective evolutionary algorithm with local search for solving the bi-objective ring star problem. Eur. J. Oper. Res. 250 (2), 377–388.

Cortés S, Gutiérrez EV, Palacio JD, Villegas JG. Districting Decisions in Home Health Care Services: Modeling and Case Study. In Workshop on Engineering Applications 2018: pp. 73-84. Springer, Cham.

Cova, T.J., Church, R.L., 2000. Contiguity constraints for single-region site search problems. Geogr. Anal. 32 (4), 306–329.

D'Amico, S.J., Wang, S.J., Batta, R., Rump, C.M., 2002. A simulated annealing approach to police district design. Comput. Oper. Res. 29 (6), 667–684.

Datta, D., Malczewski, J., Figueira, J.R., 2012. Spatial aggregation and compactness of census areas with a multiobjective genetic algorithm: a case study in Canada. Environ. Plan. B: Plan. Design 39 (2), 376–392.

Datta, D., Figueira, J., Gourtani, A., Morton, A., 2013. Optimal administrative geographies: An algorithmic approach. Socio-Econ. Plan. Sci. 47 (3), 247–257.

de Sousa SCG, Gomes C. The demand for healthcare services and resources: patterns, trends and challenges in healthcare delivery 2020.

Diglio, A., Peiró, J., Piccolo, C., Saldanha-da-Gama, F., 2021. Solutions for districting problems with chance-constrained balancing requirements. Omega in press.

Drexl, A., Haase, K., 1999. Fast approximation methods for sales force deployment. Manage. Sci. 45 (10), 1307–1323.

Duque, J.C., Church, R.L., Middleton, R.S., 2011. The p-Regions Problem. Geogr. Anal. 43 (1), 104–126.

Eroglu, D.Y., Kilic, K., 2017. A novel Hybrid Genetic Local Search Algorithm for feature selection and weighting with an application in strategic decision making in innovation management. Inf. Sci. 405, 18–32.

Fan, N., Zheng, Q.P., Pardalos, P.M., 2012. Robust optimization of graph partitioning involving interval uncertainty. Theoret. Comput. Sci. 447, 53–61.

Farahani, R.Z., 2012. Graph Theory for Operations Research and Management: Applications in Industrial Engineering: Applications in Industrial Engineering. IGI Global.

Fattahi, M., Govindan, K., 2020. Data-driven rolling horizon approach for dynamic design of supply chain distribution networks under disruption and demand uncertainty. Decision Sci.

Fattahi, M., Mahootchi, M., Mosadegh, H., Fallahi, F., 2014. A new approach for maintenance scheduling of generating units in electrical power systems based on their operational hours. Comput. Oper. Res. 50, 61–79.

Feldstein, P.J., 1966. Research on the demand for health services. Milbank Memorial Fund Quart. 44 (3), 128–165.

Galvão, L.C., Novaes, A.G., De Cursi, J.S., Souza, J.C., 2006. A multiplicatively-weighted Voronoi diagram approach to logistics districting. Comput. Oper. Res. 33 (1), 93–114.

R.S. Garfinkel G.L. Nemhauser Optimal political districting by implicit enumeration techniques Management Science 16 8 1970 B-495-B-508.

Gertler, P., Gaag, J.V.D., 1990. The willingness to pay for medical care: evidence from two developing countries. Johns Hopkins University Press.

Goldman, F., Grossman, M., 1978. The demand for pediatric care: An hedonic approach. J. Political Econ. 86 (2, Part 1), 259–280.

Gutiérrez-Gutiérrez, E.V., Vidal, C.J., 2015. A home health care districting problem in a rapid-growing city. Ingeniería y universidad. 19 (1), 87–113.

Harper, P.R., Shahani, A., Gallagher, J., Bowie, C., 2005. Planning health services with explicit geographical considerations: a stochastic location–allocation approach. Omega 33 (2), 141–152.

Haugland, D., Ho, S.C., Laporte, G., 2007. Designing delivery districts for the vehicle routing problem with stochastic demands. Eur. J. Oper. Res. 180 (3), 997–1010.

Hess, S.W., Weaver, J., Siegfeldt, H., Whelan, J., Zitlau, P., 1965. Nonpartisan political redistricting by computer. Oper. Res. 13 (6), 998–1006.

Jack, W., 1999. Principles of health economics for developing countries. World Bank Publ.

Jarrah, A.I., Bard, J.F., 2012. Large-scale pickup and delivery work area design. Comput. Oper. Res. 39 (12), 3102–3118.

Kalcsics J., Ríos-Mercado RZ. Districting Problems. In: Laporte G., Nickel S., Saldanha da Gama F. (eds) Location Science: Springer; 2019, pp. 705-743. Available at: https://doi.org/10.1007/978-3-030-32177-2_25.

Kalcsics, J., Nickel, S., Schröder, M., 2005. Towards a unified territorial design approach – Applications, algorithms and GIS integration. Top 13 (1), 1–56.

Karimi, F., 2019. Estimating the health services demand function, based on the population age pyramid. J. Health Based Res. 4 (4), 411–420.

Keyvanshokooh E, Kazemian P, Fattahi M, Van Oyen MP. Coordinated and Priority-based Surgical Care: An Integrated Distributionally Robust Stochastic Optimization Approach. Available at SSRN. 2020 Oct 1.

Keyvanshokooh, E., Shi, C., Van Oyen, M.P., 2021. Online advance scheduling with overtime: A primal-dual approach. Manufact. Service Oper. Manage. 23 (1), 246–266.

Kim, M.J., 2018. Multiobjective spanning tree based optimization model to political redistricting. Spatial Inform. Res. 26 (3), 317–325.

King, D.M., Jacobson, S.H., Sewell, E.C., Cho, W.K.T., 2012. Geo-graphs: an efficient model for enforcing contiguity and hole constraints in planar graph partitioning. Oper. Res. 60 (5), 1213–1228.

Konur, D., Geunes, J., 2019. Integrated districting, fleet composition, and inventory planning for a multi-retailer distribution system. Ann. Oper. Res. 273 (1–2), 527–559.

Lei, H., Laporte, G., Guo, B., 2012. Districting for routing with stochastic customers. EURO J. Transp. Logist. 1 (1–2), 67–85.

Lei, H., Laporte, G., Liu, Y., Zhang, T., 2015. Dynamic design of sales territories. Comput. Oper. Res. 56, 84–92.

Lei, H., Wang, R., Laporte, G., 2016. Solving a multi-objective dynamic stochastic districting and routing problem with a co-evolutionary algorithm. Comput. Oper. Res. 67, 12–24.

Lin M, Chin KS, Ma L, Tsui KL. A comprehensive multi-objective mixed integer nonlinear programming model for an integrated elderly care service districting problem. Annals of Operations Research, in press. Doi: https://doi.org/10.1007/s10479-018-3078-6.

Lin, H.Y., Kao, J.J., 2009. Subregion districting analysis for municipal solid waste collection privatization. J. Air Waste Manag. Assoc. 58 (1), 104–111.

Mayorga, M.E., Bandara, D., McLay, L.A., 2013. Districting and dispatching policies for emergency medical service systems to improve patient survival. IIE Trans. Healthcare Syst. Eng. 3 (1), 39–56.

Minciardi, R., Puliafito, P., Zoppoli, R., 1981. A districting procedure for social organizations. Eur. J. Oper. Res. 8 (1), 47–57.

Mitchell, M., 1998. An introduction to genetic algorithms. MIT press.

Montgomery, D.C., 2001. Design and Analysis of Experiments. John Wiley & Sons, New York.

Muyldermans, L., Cattrysse, D., Van Oudheusden, D., Lotan, T., 2002. Districting for salt spreading operations. Eur. J. Oper. Res. 139 (3), 521–532.

Nasir, J.A., Hussain, S., Dang, C., 2018. An integrated planning approach towards home health care, telehealth and patients group based care. J. Network Comput. Appl. 117, 30–41.

Pezzella, F., Bonanno, R., Nicoletti, B., 1981. A system approach to the optimal health-care districting. Eur. J. Oper. Res. 8 (2), 139–146.

Ricca, F., Simeone, B., 2008. Local search algorithms for political districting. Eur. J. Oper. Res. 189 (3), 1409–1426.

Ricca, F., Scozzari, A., Simeone, B., 2013. Political districting: from classical models to recent approaches. Ann. Oper. Res. 204 (1), 271–299.

Ríos-Mercado, R.Z., 2020. editor. Optimal districting and territory design. Springer Nature.

Ríos-Mercado, R.Z., Bard, J.F., 2019. An exact algorithm for designing optimal districts in the collection of waste electric and electronic equipment through an improved reformulation. Eur. J. Oper. Res. 276 (1), 259–271.

Ríos-Mercado, R.Z., Fernández, E., 2009. A reactive GRASP for a commercial territory design problem with multiple balancing requirements. Comput. Oper. Res. 36 (3), 755–776.

Ríos-Mercado, R.Z., López-Pérez, J.F., 2013. Commercial territory design planning with realignment and disjoint assignment requirements. Omega 41 (3), 525–535.

Ríos-Mercado, R.Z., Álvarez-Socarrás, A.M., Castrillón, A., López-Locés, M.C., 2021. A location-allocation-improvement heuristic for districting with multiple-activity balancing constraints and p-median-based dispersion minimization. Comput. Oper. Res. 126, 105106.

Salazar-Aguilar, M.A., Ríos-Mercado, R.Z., Cabrera-Ríos, M., 2011. New models for commercial territory design. Networks Spatial Econ. 11 (3), 487–507.

Salazar-Aguilar, M.A., Ríos-Mercado, R.Z., González-Velarde, J.L., 2013. GRASP strategies for a bi-objective commercial territory design problem. J. Heurist. 1–22.

Sandoval, M.G., Díaz, J.A., Ríos-Mercado, R.Z., 2020. An improved exact algorithm for a territory design problem with p-center-based dispersion minimization. Expert Syst. Appl. 146, 113–150.

Sandoval MG, Álvarez-Miranda E, Pereira J, Ríos-Mercado RZ, Díaz JA. A novel districting design approach for on-time last-mile delivery: an application on an express postal company. Preprint PISIS-RR21-02, Graduate Program in Systems Engineering, UANL, San Nicolás de los Garza, México, April 2021.

Shirabe, T., 2005. A model of contiguity for spatial unit allocation. Geograph. Anal. 37 (1), 2–16.

Shirabe, T., 2009. Districting modeling with exact contiguity constraints. Environ. Plan. B: Plan. Design 36 (6), 1053–1066.

Sopranzetti, B.J., 2015. Hedonic regression models. Handbook Financial Econ. Stat. 2119–2134.

Steiner, M.T.A., Datta, D., Neto, P.J.S., Scarpin, C.T., Figueira, J.R., 2015. Multi-objective optimization in partitioning the healthcare system of Parana State in Brazil. Omega 52, 53–64.

G. Taguchi Introduction to quality engineering: designing quality into products and processes 1986.

Wicks, E.M., Reasor, R.J., 1999. Designing cellular manufacturing systems with dynamic part populations. IIE Trans. 31 (1), 11–20.

Wu, B.Y., Chao, K.M., 2006. Shortest-paths trees. CRC Press, Spanning Trees and Optimization Problems, pp. 23–40.

Yamada, T., 2009. A mini–max spanning forest approach to the political districting problem. Int. J. Syst. Sci. 40 (5), 471–477.

Yanık, S., Bozkaya, B., 2020. A review of districting problems in health care. Opt. District. Territory Design. 31–55.

Yanık, S., Kalcsics, J., Nickel, S., Bozkaya, B., 2019. A multi-period multi-criteria districting problem applied to primary care scheme with gradual assignment. Int. Trans. Operat. Res. 26 (5), 1676–1697.

Zarandi, M.F., Mosadegh, H., Fattahi, M., 2013. Two-machine robotic cell scheduling problem with sequence-dependent setup times. Comput. Oper. Res. 40 (5), 1420–1434.

Zhong, H., Hall, R.W., Dessouky, M., 2007. Territory planning and vehicle dispatching with driver learning. Transp. Sci. 41 (1), 74–89.

Zoltners, A.A., Sinha, P., 1983. Sales territory alignment: A review and model. Manage. Sci. 29 (11), 1237–1256.