Healthcare Districting Optimization Using Gray Wolf Optimizer and Antlion Optimizer Algorithms (Case Study:South Khorasan Healthcare System in Iran)

Hiwa Farughi [a,*] , Sobhan Mostafayi[a] , Jamal Arkat [a]

[a]Department of Industrial Engineering, University of Kurdistan, Sanandaj, Iran

Abstract

In this paper, the problem of population districting in the health system of South Khorasan province has been investigated in the form of an optimization problem. Now that the districting problem is considered as a strategic matter, it is vital to obtain efficient solutions in order to implement in the system. Therefore in this study two meta-heuristic algorithms, Ant Lion Optimizer (ALO) and Grey Wolf Optimizer (GWO), have been applied to solve the problem in the dimensions of the real world. The objective function of the problem is to maximize the population balance in each district. Problem constraints include unique assignment as well as non-existent allocation of abnormalities. Abnormal allocation means compactness, lack of contiguous, and absence of holes in the districts. According to the obtained results, GWO has a higher level of performance than the ALO. The results of this problem can be applied as a useful scientific tool for districting in other organizations and fields of application.

Keywords: Healthcare system; Districting problem; Ant Lion Optimizer; Grey Wolf Optimizer.

1.   Introduction

In a general definition, the problem of districting is in the class of set covering problems, meaning grouping of basic units into districts. The districts in optimal mode should have features such as balance (in population size, people over 65 years old, unemployment rate and similar criteria), contiguous, compactness, and absence of holes (Datta et al., 2013). The concept of contiguous means moving to all points of a district without getting out from it. A compact district has a circle shaped without any curved shapes. Ultimately, if no district is in another district, the condition of absence of holes is fulfilled.

In order to solve the districting problem in real-world applications, it is necessary to show connections between regions in a network structure. This structure is an undirected graph G=(V, E) in which V is the set of vertices (cities or population regions) and E is the set connection paths between the vertices that we call them edges. More precisely, we assume that $V = \{v_i: i = 1,2, ..., |V|\}$ is the set of graph vertices with size | V |. In this set, each vertex v_i is represented by two vertical and horizontal vector properties (x_i, y_i), which show the latitude and longitude of the region. Moreover, $E = \{e_{ij}; i, j = 1,2, ..., |V|; i \neq j\}$ is the set of graph edges with size | E |. The connection between v_i and v_j is represented by e_{ij}. So any problem of districting the population regions can be equivalently considered as the problem of the graph districting. Determining the geographic limits and scope of activity are

decisions of organizations. However, it is not analyzed and investigated well in the framework of healthcare system optimization, and the districting of population areas in the health system is one of the newest applications of this problem (Datta et al., 2013).

It is noteworthy that there has been very little research on the application of districting in the healthcare system. Some of the most important works done in this area is described in the following section.

One of the fundamental researches in the field of the districting the healthcare system is the research carried out by (Ghiggi et al., 1975). The assumptions considered in this research are:

- Each region is composed of a certain number of indissoluble communities with a centralized population.
- Each community should be assigned to a district and no overlapping between districts is allowed.
- The location and capacity of each facility is specified.
- Each district must be connected in terms of geography.
- The population of each district should be within a certain range.
- The number of districts is known.

Finally, the purpose of this study was to present a hybrid method for districting territories, so that each district can be self-sufficient in terms of health systems, and ultimately the satisfaction of planners and citizens is meet.

(Minciardi et al., 1981) pointed out the decomposition of a geographic region into an indefinite number of non-

*Corresponding author Email address: h.farughi@uok.ac.ir

overlapping districts in their article. They introduced two innovative methods to achieve dense background areas to reduce computational trends in the production of justified districts. Also, the fuzzy set approach was used to rank the districts. Finally, they presented a case study for the districting of the health system.

(Pezzella et al., 1981) presented a method for zoning geographies in a region, in which optimal allocation of health services is available. The proposed method consists of two stages. In the first stage, the optimal hospital districts are determined by a mathematical programming model according to demand and capacity in the number of beds in the hospital. In the second stage, the health districts are determined by the accumulation of hospital districts with regard to the upper and lower limits of population in each district.

(Blais et al., 2003) studied the problem of districting health care in a city environment. According to this research, a territory should be divided into six districts by land division units. They also included four levels of districting: the inseparable sections of the unit, respect for the scope of the sections, connectivity, and work-balance.

(Benzarti et al., 2010) presented the problem of districting design for home health care structures based on a classification of criteria considered in different domains. Based on this research, the use of districting approach to facilitate the provision of organizations' caring services will reduce shipping costs and improve service quality. Some of the features of the districts in this study are: the inalienability of basic units, considering the administrative boundaries, the availability of the districts, the compactness, and the balance of workload.

(Benzarti et al., 2013) examined the problem of districting in the management of specific operations for dealing with home health care services. Innovation aspects of this research include providing a mixed integer programming model that considers criteria such as the separability of basic units, the balance of workload between human resources and adaptability. It should be noted that this research is the first one to provide multi-objective problem in the context of health system zoning.

(Datta et al., 2013) examined the districting of population districts in the UK's health system by designing a multi-objective problem with respect to the unique allocation constraints and contiguous. In this research, compactness criteria, attention to the delineation of other organs, population size balance, population age balance, and equilibrium in the population economic status of the districts were considered through the design of five-objective functions. At the end, a multi-objective genetic algorithm was used to solve the problem.

(Gutiérrez et al., 2015) examined the problem of districting home care for developing cities and the implications of population growth in urban areas, such as the development of marginal neighborhoods, increased contagious diseases, lack of government control and security, and lack of basic health services. In this study, the three outcomes were seamlessly studied, including geographical status, security conditions for access to basic units and the demand process for home health care services. Eventually, they introduced a mathematical programming model for this issue, based on

real data from one of the home health care institutions in Columbia.

(Steiner et al., 2015) presented a multi-objective algorithm for zoning the Parana state health network in Brazil. According to the results of the proposed model, 399 population districts in the state are divided into 83 different sectors. The objective functions included minimizing population aggregation in different regions, maximizing the diversity of services provided and, finally, minimizing the distance between parts of a sector. Although the results of the proposed model have been effective in improving the regional health system, the fact is that the same diversity of population in the regions cannot be considered a perfectly adequate criterion, because in some areas for reasons such as high age, high numbers Pregnant mothers, etc., a higher level of service is needed. In addition, they have only addressed the same division of the areas, while health systems in most countries in the world provide services to different categories (urban health center, rural health center, health home).

As already stated, the problem of districting in other areas has many applications which include political districting, business districting, distribution districting. In order to solve these problems, there are also many methods, which of course, are major in the research based on the meta-algorithms. To better investigate this field research, Table 1 contains studies on the districting of the graph based on their solving methods.

Regarding the literature review, it can be seen that the districting of the health system is one of the newest applications of districting. The appropriate division of the population into the health system can greatly affect the level of performance and its mode of action (Steiner et al., 2015).

According to the studies, no research has been done on system districting in South Khorasan province yet. According to the literature review, the population density is considered as the districting criterion. In order to solve this problem, we propose two algorithms, ALO and GWO, and we will compare the results of these two algorithms too. The results of this research can be made available to health professionals in order to be able to use them as a scientific and efficient tool in their decision making.

2. Methodology

This research addresses the problem of districting of population areas with the aim of enhancing the level of public health management by improving the existing system. One of the most important factors in designing health systems is to pay attention to the balance in the population in the areas. It is important, therefore, that by balancing the population of each area, the workload between the service centers can be balanced to some extent. In the strategic plan model, it is decided that the population regions of South Khorasan province are divided into 10 main areas, and the task of monitoring the performance of health organizations in each area is to be handled by the managers of these 10 sectors. To this end, around 2000 target populations are sought and are attempted to be appropriately zoned according to the organization's intended measures. In the following, the methodology of the research is discussed.

Table 1
Research in the area of graph districting based on solving method

Method	Approach			Model type	Problem	Authors
	Meta-heuristic	heuristic	deterministic			
Simulated annealing	■			Single-objective	Police management	(D'Amico et al. 2002)
Multi criteria heuristic		■		Multi-objective	Salt spreading operation	(Muyldermans et al. 2002)
Genetic	■			Multi-objective	Electricity districting	(Bergey et al. 2003)
Tabu search	■			Single-objective	Election problem	(Bozkaya et al. 2003)
Weighted Voronoi diagram		■		Multi-objective	Delivery problem	(Galvão et al. 2006)
Genetic	■			Multi-objective	Public transportation	(Tavares-Pereira et al. 2007)
Genetic	■			Multi-objective	Land Use Management	(Datta et al. 2008)
Local search	■			Multi-objective	Election problem	(Ricca and Simeone, 2008)
Voronoi diagram		■		Multi-objective	Delivery problem	(Novaes et al. 2009)
Epsilon constraint			■	Multi-objective	Distribution problem	(Aguilar et al. 2011)
Algebra Map		■		Single-objective	Bus transportation problem	(Shirabe, 2012)
Mixed integer linear programming			■	Single-objective	The combination of facility location and network design	(Contreras et al. 2012)
Genetic	■			Single-objective	Emergency medical services location	(Knight et al. 2012)
Genetic	■			Multi-objective	healthcare	(Datta et al. 2013)
Genetic	■			Multi-objective	Determine the level of resources in surgical services	(Lin et al. 2013)
Mixed integer linear programming			■	Multi-objective	Home Health Care	(Benzarti et al. 2013)
Mixed integer linear programming			■	Single-objective	Distribution problem	(Mercado and Pérez, 2013)
Grasp	■			Multi-objective	Meter readings	(De Assis et al, 2014)
Mixed integer linear programming			■	Single-objective	Use of urban land	(Li et al, 2014)
Tabu search	■			Single-objective	Periodic vehicle routing problem	(Liu et al, 2014)
Genetic	■			Multi-objective	healthcare	(Steiner et al, 2015)

2.1. Method of research

In this section, the most important and fundamental features and characteristics of this research are presented in the framework of the categorical research methodology and design. In this regard, characteristics such as type or nature of research, research orientation, philosophical foundations or research paradigms, basic research objectives, rational research approach, research strategy, research horizons, and the main points of strength and weakness of research methodology are discussed.

2.1.1. The nature of the research method

In the common typology of research, in terms of nature and inherent nature, three types of research are presented: Quantitative research, qualitative research, and combined research (quantitative and qualitative). This type of methodology is more in the light of the overall research and integrated characteristics of the set of elements and components of research into the subject. Considering the definitions and qualities that have been proposed by quantitative and quantitative research in the books and methodological sources, it can be said that according to the philosophical paradigm governing the research as well as the goals, strategy, methods of data collection and data analysis in this research, the dominant approach in this research is quantitative.

2.1.2. Type and method of research

In general, studies are divided into three types of fundamental, developmental, and functional goals. The purpose of this research is in applied research. From another perspective, studies are divided into descriptive and experimental data according to the data collection method. The present study is descriptive in terms of collecting data. The research categorizes the data quantitatively and qualitatively in terms of data analysis.It also considers quantitative in this regard. So, in order to collect data and respond to research questions, a combination of documentary and library research has been used. A documentary study has been carried out using existing data related to the subject of library research and study using a study of constructed research.

Due to the limited resources available for management and operational decisions, all organizations need to optimize the level of resource use and their decision making. In fact, managers of organizations tend to be aware of the optimal state of their decisions and move to the point where they can reach it.There are many techniques to find the optimal points in the set of possible solutions to make decisions. One

of the most effective optimization techniques is the use of mathematical programming methods. In this case, the problem is first proposed in the form of mathematical formulation and then, with the help of existing software, an optimal solution is found. This method is nowadays considered to be the most effective management method for finding optimal solutions in a very large category of problems. In this research, we try to use this technique to optimize the problem.

Since the problem of districting is NP-hard (Steiner et al., 2015), it is necessary to use meta-heuristic algorithms in order to obtain acceptable solutions in appropriate times. Most studies on the districting of the population regions in different domains have used genetic algorithm (Kim et al., 2011). However, in this paper in order to solve the problem in large dimensions, ALO and GWO algorithms have been used. Investigating the behavior of different algorithms and comparing results can provide a wider range of information to choose the appropriate solution.

One of the most important steps in the design of meta-algorithms is to design an efficient and proficient solution representation. This subject is important because the lack of proper solution representation (a solution) will produce more number of calculations and affect the performance of the algorithm. So the following section describes the representation of the solution used in this research.

2.2. Solution representation

The chromosome is used as a vector whose number of genes is equal to the number of nodes or cities. In each gene, a real and positive number is set in a predetermined interval (for example, an interval between 0 and 100), indicating the number of the area which the gene is assigned to. In order to determine the number of districts in each gene and actually convert the real number assigned to it into an integer, one can use the following procedure.

```
1   Enter the number of districts (k) and the random number interval (Interval = [min, max]).
2   The random number generation interval is divided into (k) sub-interval of equal length.
    (Clearly, the upper bound of each interval is a multiple of (l = max−min/k).
3   Calculate for each gene (j).
4   Calculate for each (i ≤ k).
5   If the real number of the gene (j)is in the interval ([(l − 1) × i, l × i]), then the real number of
    that gene changes to the integer value (i).
6   If (i) is equal to (k), the calculation is complete; otherwise, add (i)one unit.
7   If all the genes are examined, the calculation ends and otherwise the next gene is
    examined.
```

In accordance with the above structure, the amount in all genes is converted to the integer number. For example, suppose 8 cities should be divided into 4 districts, and the random number generation interval is [0,100]. Therefore, the interval is divided into sub-intervals of length $\left(\frac{100}{4} = 25\right)$. Now, the numbers generated must be correctly analyzed in accordance with the above procedure to determine districts' number. The numbers generated in each gene are assumed to be as follows in Figure 1.

| 35.2 | 12.10 | 48.01 | 84.19 | 91.41 | 8.02 | 27.11 | 68.31 |

Fig. 1. An example of a chromosome designed to solve a problem

Now, for each gene, it is examined that the real produced number is in which sub-interval, to determine the district number. Since divisions are divided by 25 coefficients, the districts' number is as follows in each gene according to Figure 2.

| 2 | 1 | 2 | 4 | 4 | 1 | 2 | 3 |

Fig. 2. A sample of modified chromosomes

Therefore,

District 1:	City 2 and city 6
District 2:	City 1, city 3, and city 7
District 3:	City 8
District 4:	City 4 and city 5

But this solution representation cannot prevent infeasible solutions. In fact, due to the existence of a random nature in the generation of initial solutions, a modification algorithm should be used to justify solutions in the production of primary population as well as in each iteration. The structure of this algorithm is expressed as follows.

Step 1: First, all solutions are randomly constructed in accordance with the above construct.

Step 2:For each of the districts, it is checked whether the assigned members have a connection contiguous. In fact, for each of the districts, you have to get the spanning tree for the points allocated to it. In a spanning tree, all vertices have connections to each other. In this paper, Kruskal's

algorithm that is described in the attachment is used to obtain a spanning tree in each iteration.

Step 3: For each of the districts, it should be checked that the non-interconnecting nodes in that district should not occur with a single district. In fact, there should be an assurance that the internal connections of the nodes of a district occur at least with two other districts. This ensures that the district is not inside the other district.

2.3. Fitness function

After determining the solution representation, an approach must be provided to calculate the fitness function. In order to calculate the fitness of each generated solution and to continue the process of implementing the algorithms, the calculation function ofsolutions' value should be determined in such a way to create a balance in the population of the districts.Thus, in accordance with the following structure, the objective function of the problem, which in fact is thefitness function of solutions, can be calculated.

Indexes

i,j,k	City
p	District

Parameters

Inh_i	Population of city i

Decision variables

x_{ip}	takes value 1 if city i is assigned to district p, otherwise it will be zero
U_p	The total existing population of district p

$$Min\ Z = \max_{p<p'}\{|U_p - U_{p'}|\} \qquad (1)$$

s.t

$$\sum_i x_{ip}Inh_i = U_p \qquad \forall p \qquad (2)$$

$$\sum_p x_{ip} = 1 \qquad \forall i \qquad (3)$$

$$x_{ip} \in \{0,1\} \qquad (4)$$

The proposed objective function in equation (1) is minimization of the maximum variation in the population of the districts. It is worth noting that equation (2) also calculates the population in each district according to the cities assigned to it. Of course, in order to make the districting appropriate, there are several other limitations that will be explained.Equation (3) ensures that each city is allocated only to one district. Equation (4)defines the range of the decision variablex_{ip}.The structure of the algorithms is described in detail.

2.4. The GWO algorithm

might partake in hunting too. However there is no idea about the optimum location (prey) in an abstract search space. It is assumed that alpha, beta, and delta have better information about location of pray. So the first three best obtained solutions are saved and the other search agents (including the omegas) should update their positions based

GWO algorithm has been first proposed by Mirjalili in 2014. Gray wolves live in packs of 5 to 12 animals. A hierarchical community exists in each pack. Leaders of the pack are a male and a female called alphas. The alpha is responsible to make decisions concerning hunting, sleeping place, time to wake, and so on. It is significant that the alpha is not fundamentally the most powerful member but it has a good capability to manage the pack. The second level of grey wolves' hierarchy is beta that can be a male or a female. The beta follows the alpha and helps him/her in making decisions or other activities. The beta is the most suitable candidate for being the alpha as soon as the alpha passes away or becomes old. The lowest level of hierarchy is omega. The omega may seem to be a scapegoat. They have to submit to all other dominant wolves. They are the last wolves that are allowed to eat in the pack. It may seem the omega is not an important member in the pack. It is notable to mention that the pack will face difficulties in case of losing the omega. On condition that a wolf is not an alpha, beta, or omega, he/she is named delta. The delta follows alphas and betas but they dominate the omega.

2.4.1. Social hierarchy

The fittest solution is considered as alpha. The second and third best solutions are assumed to be as beta and delta, respectively. The rest of the candidate solutions are named omega. The optimization process is guided by α, β, and δ in the GWO algorithm. The ω wolves follow these three wolves.

2.4.2. Encircling prey

Grey wolves encircle the prey during the hunt. In order to model this behavior mathematically, the equations (5) to (8) are considered.

$$\vec{D} = |\vec{C}.\overrightarrow{X_p}(t) - \vec{X}(t)| \qquad (5)$$

$$\overrightarrow{X_p}(t + 1) = \overrightarrow{X_p}(t) - \vec{A}.\vec{D} \qquad (6)$$

$$\vec{A} = 2\vec{a}.\vec{r}_1 - \vec{a} \qquad (7)$$

$$\vec{C} = 2.\vec{r}_2 \qquad (8)$$

Where t shows the current iteration. $\vec{A}$ and $\vec{C}$ are coefficient vectors. $\overrightarrow{X_p}$ and $\vec{X}(t)$ are the prey's position vector and wolf's position vector, respectively. $\vec{a}$ decreases linearly from 2 to 0. $\vec{r}_1$ and $\vec{r}_2$ are random vectors in [0,1].

2.4.3. Hunting

Grey wolves can identify the prey's location and encircle it. The hunt is usually directed by alpha. Beta and delta on the position of the best search agents. Hence the equations (9) to (15) are proposed.

$$\vec{D}_\alpha = |\vec{C}_1.\vec{X}_\alpha - \vec{X}| \qquad (9)$$

$$\vec{D}_\beta = |\vec{C}_2.\vec{X}_\beta - \vec{X}| \qquad (10)$$

$$\vec{D}_\delta = |\vec{C}_3.\vec{X}_\delta - \vec{X}| \qquad (11)$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 . (\vec{D}_\alpha) \tag{12}$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 . (\vec{D}_\beta) \tag{13}$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3 . (\vec{D}_\delta) \tag{14}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{15}$$

2.4.4. Attacking prey (exploitation)

The gray wolves finish the hunt by attacking the hunt as soon as it stops. So the value of $\vec{a}$ is decreased. The fluctuation range of $\vec{A}$ also decreases. $\vec{A}$ is a random value in [-2a, 2a] where a is decreased from 2 to 0 over the course of iterations. In other words, the algorithm permits its search agents to update their position according to the location of alpha, beta, and delta.

2.4.5. Search for prey (exploration)

Grey wolves often search according to the positions of alpha, beta, and delta. They individually search for prey and attacking it together. $\vec{A}$ is utilized with random values greater than 1 or less than -1. In other words, it helps the algorithm's exploration. The $\vec{C}$ contains random values in [0, 2] and it is another element of exploration. This element provides random weights for prays to emphasize C > 1 or deemphasize C < 1 the influence of prey on defining distance in above equations. The C vector is also considered as the effect of obstacles on impending prey in nature.

2.4.6. Stopping criterion

The stopping criterion is that the number of iterations of the algorithm continues to reach a predetermined value.

2.5. The ALO algorithm

The antlions' lifecycle consists of larvae and adult phases. The lifespan occurs mostly in larvae phase to hunt. An antlion larvae digs a cone-shaped pit in sand. The larvae hides at the end of the pit and waits for prey. When the antlion becomes conscious that a prey falls to the pit, it pulls the prey under the soil by its jaw. If the prey escapes, the antlion throws sand towards the edge that causes the walls to collapse more to catch the prey.

2.5.1. Operators of ALO algorithm

A random walk is selected for modelling the prey (ants' movement) as equation (16):

$$X(t) = [0, cumsum(2r(t_1) - 1), \dots, \ cumsum(2r(t_n) - 1)] \tag{16}$$

cumsum illustrates the cumulative sum. the maximum number of iteration is n while t shows the iteration. r(t) gets 1 if the rand is greater than 0.5, and 0 otherwise. rand is a random number generated with uniform distribution in [0, 1].

Each ant's position in the optimization problem is saved in M_{Ant} matrix and its elements are the solutions variables.

Similarly, the fitness value of each ant is saved in M_{OA} matrix.

$$M_{Ant} = \begin{bmatrix} A_{1,1} & \cdots & A_{1,d} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,d} \end{bmatrix} \tag{17}$$

$$M_{ALO} = \begin{bmatrix} f([A_{1,1}, A_{1,2}, \dots, A_{1,d}]) \\ \vdots \\ f([A_{n,1}, A_{n,2}, \dots, A_{n,d}]) \end{bmatrix} \tag{18}$$

The antlions' position and fitness are also saved as equations (19) and (20).

$$M_{Antlion} = \begin{bmatrix} AL_{1,1} & \cdots & AL_{1,d} \\ \vdots & \ddots & \vdots \\ AL_{n,1} & \cdots & AL_{n,d} \end{bmatrix} \tag{19}$$

$$M_{OAL} = \begin{bmatrix} f([AL_{1,1}, AL_{1,2}, \dots, AL_{1,d}]) \\ \vdots \\ f([AL_{n,1}, AL_{n,2}, \dots, AL_{n,d}]) \end{bmatrix} \tag{20}$$

The ants' random walks are normalized as equation (21).

$$X_i^t = \frac{(X_i^t - a_i)(d_i - c_i^t)}{d_i^t - a_i} + c_i \tag{21}$$

where a_i is the minimum of random walk of i-th variable, b_i is the maximum of random walk in i-th variable, c_i^t is the minimum of i-th variable at t-th iteration, and d_i^t illustrates the maximum of i-th variable at t-th iteration. Afterwards, the antlions build pits that are the effects of random movement of the ants in the search space as the equations (22) and (23).

$$c_t^i = antlion_j^t + c^t \tag{22}$$

$$d_t^i = antlion_j^t + d^t \tag{23}$$

where c^t is the minimum of all variables in t-th iteration, c_t^i is the minimum of i-th variable in j-th iteration, d^t is the maximum of all variables in t-th iteration, and d_t^i is the maximum of i-th variable in j-th iteration.

A roulette wheel is applied to model the antlions's hunting ability. Random walks of ant inside antlion's trap is formulated by the equations (24) and (25).

$$c^t = \frac{c^t}{I} \tag{24}$$

$$d^t = \frac{d^t}{I} \tag{25}$$

where c^t is the minimum of all variables at t-th iteration, d^t is the vector including the maximum of all variables at t-th iteration, and I is a ratio. The equation (26) are proposed to catch prey and re-build the pit.

$$Antlion_j^t = Ant_i^t \quad if \ f(Ant_i^t) > f(Antlion_j^t) \tag{26}$$

where t shows the current iteration, $Antlion_j^t$ indicates the position of selected j-th antlion at t-th iteration, and Ant_i^t is the position of i-th ant at t-th iteration.

The best antlion is considered as an elite in each iteration in this algorithm. Each ant randomly moves surrounding the elite by means of the roulette wheel and the selected antlion as equation (27).

$$Ant_i^t = \frac{R_A^t + R_E^t}{2} \qquad (27)$$

where R_A^t shows the random walk around the antlion selected by the roulette wheel at t-th iteration, R_E^t is the random walk around the elite at t-th iteration, and Ant_i^t illustrates the position of i-th ant at t-th iteration.

2.5.2. Stopping criterion

The stopping criterion is that the number of iterations of the algorithm continues to reach a predetermined value.

3. Computational Results

In order to evaluate the efficiency of the proposed algorithms, a number of numerical examples are generated in accordance with the real-world conditions and the following procedure. In these examples, some population units (C) and a number of districts (D) are specified. A remarkable fact in these examples is the population and the distance between the population units that should be determined in a precise manner. For this purpose, a two-dimensional environment is considered for each example. For all of the generated examples, the dimensions of the intended area are within the range of $U(100,150)$. Then for each dimension x and y, $|C|$ non-repetitive random number is selected and considered as a set $\{i \in C | (x_i, y_i)\}$. These sets are, in fact, the geographical coordinates of population points. The relationship between points must be in the form of a connected graph, and the distance between points is calculated by means of Euclidean distance. According to the above, the problem structure is formed, and in the next step the population of each point must be determined. At this stage, firstly, the population points are divided into two categories of rural and urban areas. The ratio of the number of rural points to urban points can be determined by the following relationships. These relations are presented in accordance with current conditions in Iran. According to the latest census data in Iran, the total number of population points (T) is 97877, of which the number of villages (NV) is 96546 and the number of cities (NC) is 1331. The total population of the country (TP) is about 80 million with the portion of villages (VPP) of 28% and the city portion (CPP) of 72% of the total population. Therefore, the ratio of the number of rural areas (NVP) and urban areas (NCP) as well as the population of each point for production examples can be presented as equations (28) to (31).

$$MCP = \frac{TP \times CPP}{NC}, \qquad (28)$$

$$MVP = \frac{TP \times VPP}{NV}, \qquad (29)$$

$$NCP = \frac{NC}{T}, \qquad (30)$$

$$NVP = \frac{NV}{T}. \qquad (31)$$

Initially, $\lceil |C| \times NCP \rceil$ and $\lfloor |C| \times NVP \rfloor$ points are randomly selected among the number of population points $|C|$ and considered as the urban points and rural points, respectively. The population of each of the points is also chosen for cities and villages using the normal distribution, respectively, $N \sim (MCP, \rho)$ and $N \sim (MVP, \varphi)$, where ρ and φ are standard deviation of normal distribution and are calculated as equations (32) and (33).

$$\rho = \sqrt{\frac{\sum_{i \in city}(MCP - Inh_i)^2}{\lceil |C| \times NCP \rceil - 1}}, \qquad (32)$$

$$\varphi = \sqrt{\frac{\sum_{i \in village}(MVP - Inh_i)^2}{\lceil |C| \times NVP \rceil - 1}}. \qquad (33)$$

After generating numerical examples in accordance with the above procedure, in order to evaluate the efficiency of the proposed algorithms, we solve each of these examples by all algorithms. It should be noted that all examples are solved with the help of MATLAB software version 2016, powered by a 3.2GHz processor, random access memory of 4GB and Windows 10 operating system. Table 2 shows the results of these calculations.

According to the information in Table 2, GWO and ALO algorithms largely have similar results. This is well illustrated in the GAP computational part of Table 2 (the algorithm has better performance with (*) specified). In order to better understand the performance of algorithms on various problems, we provide the solutions separately in the form of the following graphs.

According to Figure 3, in some of the modes, the solution of GWO and ALO is completely overlapping or with very little GAP. In the case of the solution time, it can be seen that the GWO algorithm has a lower computational time. Figure 4 shows a fair comparison between solving times.

It should be noted that districting is usually the strategic level decision of the organization, and implementation time is not an important factor. In such cases, the quality of solution is much more important than run time. However, in order to evaluate the efficiency of algorithms, the solving time has also been reported.

According to the results of algorithms, the GWO algorithm has the highest performance compared to the other; therefore, the algorithm is used to solve this problem. In the following section, we describe the structure of case study, which is the health system of the southern Khorasan province.

Table 2
Comparison of the results of algorithms

| Mode	Size (city-district)	GWO		ALO		GAP
		Objective function	Time (s)	Objective function	Time (s)	GWO-ALO*
1	10-2	**17788**	42	**17788**	60	0
2	12-2	**18878**	99	**18878**	61	0
3	15-2	**13663**	109	**13663**	**79**	0
4	20-3	**18967**	121	**18967**	160	0
5	20-4	**14302**	**129**	**14302**	160	0
6	30-5	**16438**	**144**	**16438**	171	0
7	35-5	**19416**	171	**19416**	257	0
8	35-6	**16852**	**175**	**16852**	306	0
9	40-5	**19154**	206	**19154**	318	0
10	45-5	**17210**	**207**	**17210**	377	0
11	50-6	**14166**	**221**	**14166**	307	0
12	55-6	**14524**	**225**	**14524**	311	0
13	60-6	11601	**228**	**11155**	332	0.04
14	70-6	**19357**	**238**	**19357**	336	0
15	80-7	**16099**	**246**	**16099**	354	0
16	85-8	12035	**361**	**11572**	393	0.04
17	90-8	15545	**303**	**14805**	502	0.05
18	100-8	**15747**	**306**	**15747**	503	0
19	110-9	18112	**408**	**17087**	504	0.06
20	120-9	12488	**426**	**11781**	542	0.06
21	130-9	20724	**427**	**19551**	593	0.06
22	150-10	**15861**	**427**	**15861**	597	0
23	200-10	**18051**	**453**	**18051**	607	0
24	230-10	18973	**487**	**17568**	615	0.08
25	250-10	14666	**514**	**13580**	649	0.08
26	300-11	**16526**	**531**	**16526**	665	0
27	350-11	**17563**	**539**	**17563**	680	0
28	380-11	**16132**	**581**	**16132**	705	0
29	400-11	19392	**613**	**17956**	714	0.08
30	400-12	**17549**	**652**	**17549**	741	0
31	450-12	**13705**	**683**	**13705**	742	0
32	450-13	16490	**798**	**15128**	788	0.09
33	480-13	**14591**	**832**	**14591**	795	0
34	500-14	**14345**	**862**	**14345**	798	0
35	550-15	**17745**	**875**	**17745**	949	0
36	550-16	**16386**	**896**	**16386**	958	0
37	600-16	**19610**	**981**	**19610**	969	0
38	700-16	17013	**988**	**15327**	1006	0.11
39	750-16	16332	**1000**	**14582**	1012	0.12
40	800-17	19209	**1001**	**17151**	1025	0.12
41	800-18	22502	**1017**	**19913**	1033	0.13
42	900-18	15809	**1017**	**13990**	1070	0.13
43	1000-19	**12553**	**1031**	**12553**	1074	0
44	1000-20	21915	**1044**	**19224**	1077	0.14
45	1100-20	15906	**1054**	**13953**	1129	0.14
46	1200-20	20934	**1066**	**18363**	1151	0.14
47	1300-20	21870	**1071**	**19184**	1163	0.14
48	1500-20	**19643**	**1075**	**19643**	1176	0
49	1800-20	**14471**	**1090**	**14471**	1177	0
50	2000-20	16505	**1096**	**13987**	1192	0.18
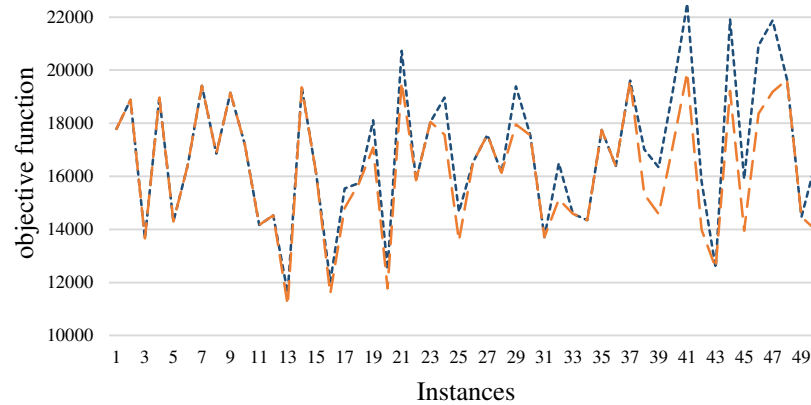
Fig. 3. Comparison of the values of the objective function in different algorithms
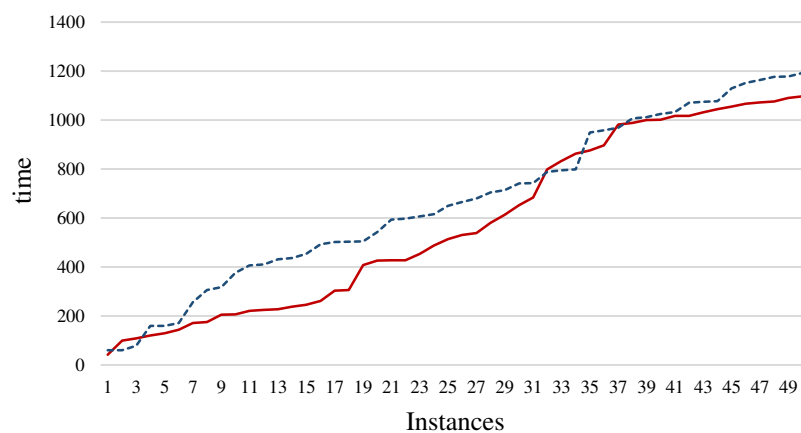

Fig. 4. Comparison of the solving time of the proposed algorithms

3.1. Case study

Managing healthcare has always been one of the biggest concerns of healthcare managers. The value of medical services in 2002 was approximately $ 24 billion, and was projected to increase to $ 31 billion by 2007. Iran, with a population of about 80 million, is one of the populous countries of the Middle East. Demographically, Iran is one of the young countries in the region, and it has always been synchronous with population growth, with many requests for various public services.The population will soon be getting enough age to form a new family. This will increase population growth considerably and rise the need for public health, and health services in particular. It is expected that all treatment costs will increase from $ 24.3 trillion in 2008 to $ 80 trillion by 2018, indicating an increase in the demand for medical services.

The total health care costs of Iran in 2005 were 4.45% of gross domestic product. 73% of Iranians are covered by health insurance. Iran has rank of 58th in level of health and the overall performance of the health system and 93rd among the world's governments according to the latest report of World Health Organization (WHO).

The health status in Iran has improved over the past two decades. Iran has been able to provide public health prevention services through the establishment of a comprehensive primary health service, resulting in a significantly lowering maternal and child mortality rate, and, interestingly, the average life span of life has increased since birth. Mortality of infant and children under the age of 5 years in 2000 compared to 1970, of which 122 out of every 1000 babies and 191 out of 1000 children under age 5 were 28.6 and 35.66 out of 1000 live births decreases, respectively.

As outlined in the previous sections, in order to implement the strategic health plan model of the South Khorasan province, the population parts of this province should be divided into 10 districts. Since the size of the population of the province cannot be considered totally because of large size, about 2000 population points are considered in this study. It should be noted that in solving districting problems, the connections between points is considered as a connected graph.

After solving the problem with the help of the GWO algorithm that showed the highest performance in the algorithm evaluation section, the resulting districting is presented in Fig. 5. It is worth noting that the following results are the best results of 10 independent implementation of the algorithm. Also, the maximum number of repetitions is 2000.
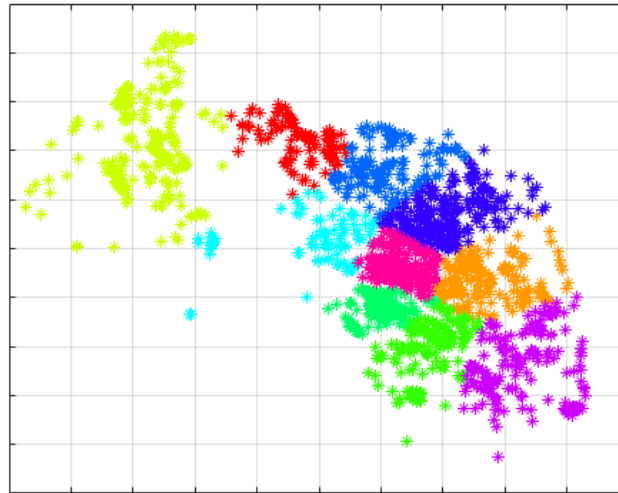
Fig. 5. Graphical structure derived from GWO algorithm

As can be seen, population units are divided into ten unique districts. The purpose of the unique assignment is that, according to the definition, each node should be assigned exactly to a district that is well-considered in the results. It is also clear that no unusual assignment has been created. In fact, with regard to the concept of contiguous, it can be seen that no district is formed into two or more parts, and all the nodes in a district are directly or indirectly related to each other. In addition, no district is enclosed within others. This limitation, which is introduced as a limitation of the absence of holes in the districts, is well described by the solution modification algorithm. Therefore, there is a guarantee that the modification algorithm always verifies the solutions in all stages of the algorithm. But another important matter is the appearance of the districts. As previously stated, in a proper districting, the structural form of all districts should be as circular as possible, which is also well considered. But in the case of the size of the districts, it should be noted that some of the districts have larger dimensions. This is because

of the difference in population density in the country demographic. In fact, in the desert areas, the distance between the points is greater, and this causes the structure of the districts in the eastern, southeast and central regions to be larger than the northern and western parts of the country. According to the fitness function used in this study, which has focused on the equilibrium in the population of the districts, the values of the final objective function and the population in each district are presented in Table 3. In addition to the population in each district, in this table, the population deviation of each district from the average population size of each district is also presented. This information can be useful for evaluating the efficiency of the algorithm, because no matter how much the departures of population density are from the mean value, it indicates that the algorithm is performing accurate calculations. It should be noted that the total population of the points considered is 724475, the average value for the 10 districts is 5372425.

Table 3
Population size and deviation percentage from the average population in each district

Number of district	District population	Percentage deviation from average population
1	91405	21
2	60937	19
3	74478	3
4	101561	29
5	54166	34
6	88020	18
7	40624	27
8	60937	19
9	84634	14
10	67713	7

As can be seen, the highest difference is 34% , which is negligibleto the size of the total population. In fact, in all districts, the size of the population is largely equal to the average population. Therefore, we can say that the proposed algorithm has a good performance.

Another important matter is the convergence of the final solution. According to the nature of meta-heuristic

algorithms that use a random approach, there is a prediction that in the initial iterations, the degree of convergence is much higher than the end ones. The reason for this is the transfer of the algorithm from the phase of diversity to the phase of interdependence. Figure 6 shows the rate of convergence to the final solution.
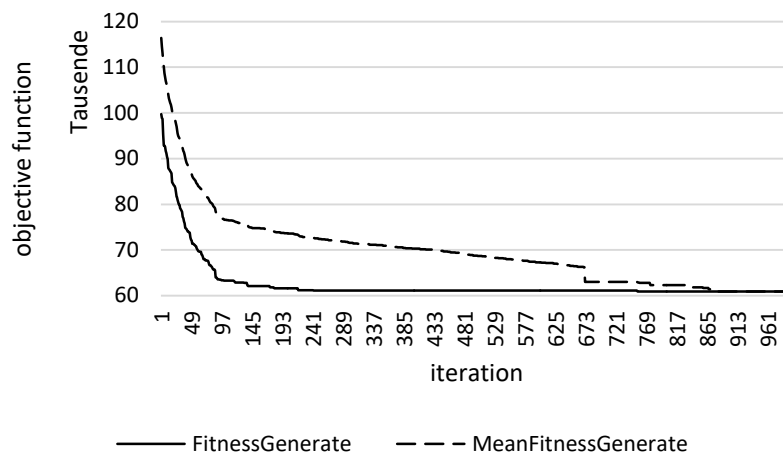
Fig. 6. Convergence to the final solution in GWO algorithm

According to the above figure, it is seen that in the initial iterations the convergence of the algorithm is very fast and the improvement in solutions is quite evident, but gradually by increasing the number of iterations and improving the value of the objective function, the algorithm goes out from the phase of diversity and enters the semanticist phase, resulting in an improvement in the solution rate to a point where no improvement has been made in repetitions of 760. Finally, the final solution has been reported.

4. Discussion and Conclusion

In this research, the problem of districting of the population units of South Khorasan province has been studied. According to the studies, the decision is to divide the health system of the South Khorasan Province into 10 health districts, so that each districtcan do the task of overseeing the performance of its health organizations. For this purpose, around 2000 population points have been considered. Attempts have been made to modify these population units in accordance with the criteria of the organization. One of the most important criteria for districting can be to maximize the population balance in each district and minimize the maximum distance between points in each district. The objective of the proposed problem is to create a population balance in the formed districts. This goal is achieved by minimizing the maximum difference between population of districts. Since the problem of districting the population units is NP-hard, meta-heuristic algorithms are used to obtain appropriate results in real-world problems. In this paper, the GWO and ALO algorithms are presented. But the key to using these algorithms is to ensure their correct performance. To do this, first, a number of random examples are generated based on real cases and the results of the algorithms are compared to each other. According to these results, the GWO algorithm has a higher level of performance than the ALO algorithm. In the end, according to available information, the population of South Khorasan province is divided into 10 districts by means of algorithms. In order to carry out further research to expand the dimension of the problem, it is suggested to apply new meta-heuristic algorithms and compare the results with the results of the existing algorithms.

Appendixes

The structure of Kruskal's algorithm is described as below.

1.	functionKruskal(G)
2.	**for**eachvertexvin**G**do
3.	DefineanelementaryclusterC(v)←{v}.
4.	InitializeapriorityqueueQtocontainalledgesinG,**using**theweightsaskeys.
5.	DefineatreeT←Ø//T will ultimately contain the edges of the MST
6.	// n is total number of vertices
7.	**while**Thasfewerthann-1edges**do**
8.	// edge u,v is the minimum weighted route from/to v
9.	(u,v)←Q.removeMin()
10.	// prevent cycles in T. add u,v only if T does not already contain a path between u and v.
11.	// Note that the cluster contains more than one vertex only if an edge containing a pair of
12.	// the vertices has been added to the tree.
13.	LetC(v)betheclustercontainingv,andletC(u)betheclustercontainingu.
14.	**if**C(v)≠C(u)then
15.	Addedge(v,u)toT.
16.	MergeC(v)andC(u)intoonecluster,thatis,**union**C(v)andC(u).
17.	**return**treeT

References

Benzarti, E., Sahin, E., & Dallery, Y. (2010). Modelling approaches for the home health care districting problem. *In 8th International Conference of Modeling and Simulation-MOSIM*, 10-12

Benzarti, E., Sahin, E., & Dallery, Y. (2013). Operations management applied to home care services: Analysis of the districting problem. *Decision Support Systems*, 55(2), 587-598.

Bergey, P. K., Ragsdale, C. T., & Hoskote, M. (2003). A simulated annealing genetic algorithm for the electrical power districting problem. *Annals of Operations Research*, 121(1-4), 33-55.

Blais, M., Lapierre, S. D., & Laporte, G. (2003). Solving a home-care districting problem in an urban setting. *Journal of the Operational Research Society*, 54(11), 1141-1147.

Bozkaya, B., Erkut, E., & Laporte, G. (2003). A tabu search heuristic and adaptive memory procedure for political districting. *European Journal of Operational Research*, 144(1), 12-26.

Contreras, I., Fernández, E., & Reinelt, G. (2012). Minimizing the maximum travel time in a combined model of facility location and network design. Omega, 40(6), 847-860.

D'Amico, S. J., Wang, S. J., Batta, R., & Rump, C. M. (2002). A simulated annealing approach to police district design. *Computers & Operations Research*, 29(6), 667-684.

Datta, D., Figueira, J. R., Gourtani, A. M., & Morton, A. (2013). Optimal administrative geographies: an algorithmic approach. *Socio-Economic Planning Sciences,* 47(3), 247-257.

Datta, D., Fonseca, C. M., & Deb, K. (2008). A multi-objective evolutionary algorithm to exploit the similarities of resource allocation problems. *Journal of Scheduling,* 11(6), 405.

De Assis, L. S., Franca, P. M., & Usberti, F. L. (2014). A redistricting problem applied to meter reading in power distribution networks. *Computers & Operations Research*, 41, 65-75.

Galvão, L. C., Novaes, A. G., De Cursi, J. S., & Souza, J. C. (2006). A multiplicatively-weighted Voronoi diagram approach to logistics districting. *Computers & Operations Research,* 33(1), 93-114.

Ghiggi, C., Puliafito, P. P., & Zoppoli, R. (1975, September). A combinatorial method for health-care districting. In IFIP Technical Conference on Optimization Techniques, 116-130. Springer, Berlin, Heidelberg.

Grefenstette, J. J. (Ed.). (2013). Genetic algorithms and their applications: proceedings of the second international conference on genetic algorithms. Psychology Press.

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.

Kim, J., Hwang, I., Kim, Y. H., & Moon, B. R. (2011, July). Genetic approaches for graph partitioning: a survey. *In Proceedings of the 13th annual conference on Genetic and evolutionary computation,* 473-480. ACM.

Knight, V. A., Harper, P. R., & Smith, L. (2012). Ambulance allocation for maximal survival with heterogeneous outcome measures. Omega, 40(6), 918-926.

Li, W., Church, R. L., & Goodchild, M. F. (2014). An extendable heuristic framework to solve the p-compact-regions problem for urban economic modeling. Computers, *Environment and Urban Systems*, 43, 1-13.

Lin, R. C., Sir, M. Y., & Pasupathy, K. S. (2013). Multi-objective simulation optimization using data envelopment analysis and genetic algorithm: Specific application to determining optimal resource levels in surgical services. Omega, 41(5), 881-892.

Liu, R., Xie, X., & Garaix, T. (2014). Hybridization of tabu search with feasible and infeasible local searches for periodic home health care logistics. Omega, 47, 17-32.

Minciardi, R., Puliafito, P. P., & Zoppoli, R. (1981). A districting procedure for social organizations. *European Journal of Operational Research*, 8(1), 47-57.

Muyldermans, L., Cattrysse, D., Van Oudheusden, D., & Lotan, T. (2002). Districting for salt spreading operations. *European Journal of Operational Research*, 139(3), 521-532.

Novaes, A. G., de Cursi, J. S., da Silva, A. C., & Souza, J. C. (2009). Solving continuous location–districting problems with Voronoi diagrams. *Computers & Operations Research,* 36(1), 40-59.

Pezzella, F., Bonanno, R., & Nicoletti, B. (1981). System Approach to the Optimal Health-Care Districting. *Europ. J. Oper. Res*., 8(2), 139-146.

Ricca, F., & Simeone, B. (2008). Local search algorithms for political districting. *European Journal of Operational Research*, 189(3), 1409-1426.

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

Salazar-Aguilar, M. A., Ríos-Mercado, R. Z., & González-Velarde, J. L. (2011). A bi-objective programming model for designing compact and balanced territories in commercial districting. Transportation Research Part C: Emerging Technologies, 19(5), 885-895.

Shirabe, T. (2012). Prescriptive modeling with map algebra for multi-zone allocation with size constraints. Computers, *Environment and Urban Systems*, 36(5), 456-469.

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.

Storn, R., & Price, K. (1997). Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization, 11(4), 341-359.

Tavares-Pereira, F., Figueira, J. R., Mousseau, V., & Roy, B. (2007). Multiple criteria districting problems. *Annals of Operations Research*, 154(1), 69-92.