A queuing location-allocation model for a capacitated health care system

M Pouraliakbarimamaghani, M. Mohammadi, A Mirzazadeh

Scientia Iranica E (2017) 24(2), 751–764

SCIENTIA
IRANICA

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
www.scientiairanica.com

A queuing location-allocation model for a capacitated health care system

M. Pouraliakbarimamaghani, M. Mohammadi* and A. Mirzazadeh

Department of Industrial Engineering, Faculty of Engineering, Kharazmi University, Tehran, Iran.

Received 3 January 2015; received in revised form 13 January 2016; accepted 12 March 2016

KEYWORDS
Health care system;
Mixed-integer programming;
Queuing theory;
Capacitated system;
Genetic algorithm;
Simulated annealing algorithm;
Hybrid algorithm.

Abstract. The aim of the present paper is to propose a location-allocation model for a capacitated health care system. This paper develops a discrete modeling framework to determine the optimal number of facilities among candidates and optimal allocations of the existing customers to operating health centers in a coverage distance. In doing so, the total sum of customer and operating facility costs is minimized. Our goal is to create a model that is more practical in the real world. Therefore, setup costs of hospitals are based on the costs of customers, fixed costs of establishing health centers, and costs based on the available resources in each level of hospitals. In this paper, the idea of hierarchical structure has been used. There are two levels of service in hospitals, including low and high levels, and sections at different levels that provide different types of services. The patients refer to different sections of the hospital according to their requirements. To solve the model, two meta-heuristic algorithms, including genetic and simulated annealing algorithms and their combination, are proposed. To evaluate the performance of the three algorithms, some numerical examples are produced and analyzed using the statistical test in order to determine which algorithm works better.

1. Introduction

Many economical decision problems concern selecting and placing certain facilities to serve given demands efficiently, so the location-allocation of facilities in service systems is an essential factor of their performance. These models are classified into discrete and continues models. Current et al. [1] carried out a study on discrete Location-Allocation (LA) models; a discrete location-allocation model has been studied in this paper. One important situation less addressed in the literature review is to balance the service among customers, in addition to minimizing the cost of establishing the facilities. This issue is of crucial importance, especially in the public sector. Therefore, we have developed a model that considers customer satisfaction and the setup costs together. The facility location problem, introduced by Balinski [2], addressed the problem of locating a new set of facilities in a way that minimizes the sum of those two costs. One important topic not mentioned in the health care's literature review is the particular consideration of the costs of installation and operation in hospitals. In this model, this cost will be discussed in detail. So, we want our model to be more practical in the real world; we represent our establishment costs according to the three categories of costs: the patient costs, the fixed costs of establishing health centers, and finally, the costs based on the available resources in each level of hospitals, respectively. In this paper, the idea of hierarchical structure has been used. According to the literature of hierarchical structure, there exist many

*. *Corresponding author. Tel.: +98 21 88830891; Fax: +98 21 88329213*
E-mail addresses: mahsap58@yahoo.com (M. Pouraliakbarimamaghani); Mohammadi@khu.ac.ir (M. Mohammadi); a.mirzazadeh@aut.ac.ir (A. Mirzazadeh)

hierarchical structures at service networks, such as health care systems. In these systems, general centers provide low-level services, such as primary health care, and specialized hospitals provide high-level services. In this research, we use this structure inside any systems, because there are some low and high levels and sections at different levels of hospitals that provide different types of services and have different resources. So, the other goal of the model is to determine the optimal capacity of resources at different sections of the two levels of the hospitals according to the requirements of patients. The constraints of hospital resource capacity have been considered to avoid the shortages of different resources in different parts (low and high levels) of the hospitals.

2. Literature review

LA problems have been studied by many researchers since Weber [3] has proposed it for the first time. Various facility location problems are formulated and studied for decades, including those based on p-median and fixed charge facility location problems and their extensions [4-6]. The Location Set Covering Problem (LSCP) was developed by Toregas et al. [7]. Facility location-allocation problems in many practical settings arise from various industries, including health care and emergency services to manufacturing networks [8-11]. The first facility location model for health care systems was introduced by Hakimi [12]. Also, it was followed by many innovative efforts of the other researchers. Guerriero et al. [13] studied the problem of location and reorganized the Calabrian health care network. Ghaderi et al. [14] presented a model that is concerned with the determination of the optimal locations of incapacitated health care systems. Vatsa et al. [15] studied a multi-period problem of allocating doctors to primary health centers. Baray et al. [16] presented a hierarchical location-allocation model with the objective of combining a maximum covering model and p-center models to optimize locations of the three levels of maternity hospitals found in France. Mohammadi et al. [17] proposed a bi-objective reliable location model for health care management under limited capacity, and a patient queue system with two patient groups was created. Cooper [18] categorized the LA problems into two different classes: One of them is called uncapacitated LA problem, considered by many researchers such as Damgacioglu et al. and Kratica et al. [19,20]. The other category of LA problems is considered by many researchers, including Alizadeh et al., Zhou and Liu, and Marinakis [21-23]. Comprehensive review of hierarchical location models can be found in the studies of Sahin and Sural and Zanjirani Farahani et al. [24,25]. Moreover, The first fuzzy model for location-allocation in the hierarchical systems was developed by Shavandi and Mahlooji et al. [26]. They introduced a fuzzy hierarchical queuing location-allocation model for MCLP in coherent systems. Moreover, in another work, Shavandi and Mahalooji [27] developed fuzzy hierarchical queuing models for MCLP in both nested and refferal systems. Furthermore, a successively inclusive hierarchical model for location of health centers in terms of the transfer of patients from a lower to higher level of health centers was studied by Alinaghian et al. [28].

3. Problem definition

We intend to answer the questions of how many facilities should be opened, where the opened facilities in the candidate sites should be located, and eventually what demand nodes each opened facility should cover. In so doing, the total cost, including transportation and fixed location and opening costs of the facilities, is minimized.

Assumptions

In order to model the problem, the paper considers the following assumptions, indices, parameters, and variables:

1. A discrete location-allocation model in health care systems is discussed in this paper;
2. Each facility acts as $M/M/C/K$ queuing system; there are C servers in each facility, and the capacity of each facility is finite (K);
3. The health care facilities have to cover the total demand in coverage distance (D).

Indices

i Index for customer nodes ($i = 1, \cdots, M$, where M is the number of customer nodes);

j Index for potential service nodes ($j = 1, \cdots, N$, where N is the number of potential service nodes);

k Index for low-level sections inside the hospitals ($k = 1, \cdots, K$);

z Index for high-level sections inside the hospitals ($z = 1, \cdots, Z$);

r_j^L Index for resource types inside the low-level sections of the hospital j ($r_j^L = 1, \cdots, R_j^L$);

r_j^H Index for resource types inside the high-level sections of the hospital j ($r_j^H = 1, \cdots, R_j^H$).

Parameters

v_1, v_2 The weights of two objective functions;

c_j^S Setup cost for facility j;

c_j^F Fixed cost of opening a facility at node j such as land costs, labor costs, equipment costs, and so on;

a_i The population in demand node i;

$d_{i,j}$ Distance matrix from demand node i to facility node j;

D Maximum distance that a service node can serve a customer node (maximum coverage distance), when $d_{i,j}$ is smaller than or equal to the maximum coverage distance; then the allocation variable can be 1; otherwise, it is zero. $N_i = \{j/d_{i,j} \leq D\}$;

U^C Upper bound for hospital capacity;

U^L Upper bound for resource capacity in low levels;

U^H Upper bound for resource capacity in high levels;

$w_{k,j}$ Indicates whether or not the arrival patients to the hospitals need low-level section k of hospital j (binary parameter);

$w'_{z,j}$ Indicates whether or not the low-level patients need high level section z of hospital j (binary parameter);

$w''_{z,j}$ Indicates whether or not the arrival patients to the hospitals need high-level section z of hospital j (binary parameter), (for patients who are allocated to the high-level sections directly without going to the low-level sections);

$p_{k,j}$ Refers to the fraction of arrival patients to hospital j that needs low-level section k (a number between (0, 1));

$p'_{z,j}$ Refers to the fraction of arrival patients to low-level section k that needs high-level section z (a number between (0, 1));

$p''_{z,j}$ Refers to the fraction of arrival patients to hospital j that needs high-level section z (a number between (0, 1)), (for patients that directly visit the high-level sections without going to the low-level sections);

$y'_{k,j}$ Location parameter shows that low-level section k at hospital j is open;

$y''_{z,j}$ Location parameter shows that high-level section z at hospital j is open;

λ_j Patient arrival rate at open facility j;

$\lambda'_{k,j}$ Patient arrival rate at low-level section k of hospital j;

$\lambda''_{z,j}$ Patient arrival rate at high-level section z of hospital j.

Variables

$x_{i,j}$ Allocation variable that takes value 1 if population at demand node i is allocated to facility location j, and otherwise is zero;

y_j Location variable that takes value 1 if facility is located at node j, and zero otherwise;

cap_j Capacity of facility j which is opened;

$x'_{k,z,j}$ Allocation variable that takes value 1 if patients at hospital j are allocated to low-level section k and are referred to high level section z then; otherwise, it is zero;

$\text{Rcap}_{r_j^L,k,j}$ Resource capacity at low-level section k of hospital j;

$\text{Rcap}'_{r_j^H,z,j}$ Resource capacity at high-level section z of hospital j.

3.1. Mathematical model

Our goal is to create a model that is more practical in the real world. Therefore, one of the setup costs of the hospitals is based on the costs of clients. In this case, the average costs of all customers from each facility are considered as maximum setup costs of hospitals. It is all for the sake of considering the clients' welfare. So average transportation costs of all clients to each hospital (average distance multiplied by the average demands of all customers from each hospital) are considered as a base cost of building hospitals to consider the welfare of the clients and to make the model closer to the reality; the mathematical model is developed as follows:

$$\text{Minimize } Z = v_1 \sum_{j=1}^{N} \sum_{i=1}^{M} a_i d_{i,j} x_{i,j} + v_2 \sum_{j=1}^{N} \left(c_j^S + c_j^F + c_j^C\right) y_j, \tag{I}$$

$$\sum_{j=1}^{N} y_j \leq N, \tag{1}$$

$$x_{i,j} \leq y_j \qquad \forall\, i \in M, \qquad j \in N, \tag{2}$$

$$\sum_{j\in N_i} x_{i,j} = 1 \quad \forall\, i \in M, \quad N_i = \{j/d_{i,j} \le D\}, \tag{3}$$

$$\lambda_j = \sum_{i=1}^{M} a_i x_{i,j} \qquad \forall\, j \in N, \tag{4}$$

$$\lambda_j \le \mathrm{cap}_j \qquad \forall\, j \in N, \tag{5}$$

$$\mathrm{cap}_j \le U^C y_j \qquad \forall\, j \in N, \tag{6}$$

$$\lambda'_{k,j} = \lambda_j w_{k,j} p_{k,j} \qquad \forall\, k \in K, \qquad j \in N, \tag{7}$$

$$\lambda''_{z,j} = \sum_{k=1}^{K} \left[\lambda'_{k,j} x'_{k,z,j} w'_{z,j} p'_{z,j}\right] + \left[\left(\lambda_j - \sum_{k=1}^{K} \lambda'_{k,j}\right).(w''_{z,j} p''_{z,j})\right], \quad \forall\, z \in Z, \qquad j \in N, \tag{8}$$

$$\lambda'_{k,j} \le \mathrm{Rcap}_{r_j^L,k,j} \quad \forall\, k \in K, \qquad j \in N, \qquad r_j^L \in R_j^L, \tag{9}$$

$$\lambda''_{z,j} \le \mathrm{Rcap}'_{r_j^H,z,j} \quad \forall\, z \in Z, \qquad j \in N, \qquad r_j^H \in R_j^H, \tag{10}$$

$$\mathrm{Rcap}_{r_j^L,k,j} \le U^L y'_{k,j} \quad \forall\, r_j^L \in R_j^L, \qquad k \in K, \qquad j \in N, \tag{11}$$

$$\mathrm{Rcap}'_{r_j^h,z,j} \le U^H y''_{z,j} \quad \forall\, r_j^H \in R_j^H, \qquad z \in Z, \qquad j \in N, \tag{12}$$

$$\sum_{k=1}^{K} \mathrm{Rcap}_{r_j^L,k,j} = \mathrm{cap}_j, \quad \forall\, r_j^L \in R_j^L, \quad j \in N, \tag{13}$$

$$\sum_{z=1}^{Z} \mathrm{Rcap}'_{r_j^H,z,j} = \mathrm{cap}_j, \quad \forall\, r_j^H \in R_j^H, \quad j \in N, \tag{14}$$

$$c_j^{C'} = \sum_{r_j^L \in R_j^L} \sum_{k=1}^{K} \mathrm{Rcap}_{r_j^L,k,j} + \sum_{r_j^H \in R_j^H} \sum_{z=1}^{Z} \mathrm{Rcap}'_{r_j^H,z,j} \quad \forall\, j \in N, \tag{15}$$

$$x_{i,j}, y_j, x'_{k,z,j} \in \{0,1\} \quad \forall\, i \in M, \qquad k \in K, \qquad z \in Z, \qquad j \in N, \tag{16}$$

$$\mathrm{cap}_j \ge 0, \quad \mathrm{Rcap}_{r_j^L,k,j} \ge 0, \quad \mathrm{Rcap}'_{r_j^h,z,j} \ge 0, \text{integer} \quad \forall\, r_j^L \in R_j^L, \qquad r_j^H \in R_j^H, \qquad k \in K, \qquad z \in Z, \qquad j \in N. \tag{17}$$

Our objective function is to balance facility setup and patient costs by considering weights for objective functions. It will allow decision-makers to decide based on customer convenience or startup fees. Constraint 1 sets an upper bound for the maximum number of the opened facilities. Constraint 2 assumes that allocation variable can take value 1 only when a facility has already been located at node j. Constraint 3 states that customer demand is satisfied by only one opened facility and all demand nodes should be covered in the coverage distance, $N_i = \{j/d_{i,j} \le D\}$. Constraint 4 calculates the patient arrival rate at open facility j. Constraint 5 assures that the input of each facility should be less than its capacity. Constraint 6 assures that system capacity at open facility j is less than upper bound (this shows the system capacity is finite). In order to discuss Constraints 7 and 8, the following assumptions have been considered. In this model, it is assumed that each hospital consists of both low and high levels. Low-level sections of the hospitals offer low-level services, and high-level sections of the hospitals offer high-level services. Each of these two levels has different sections. The low-level sections of the hospitals consist of sections, such as hospital emergency departments, and general practitioners departments; the high-level sections of the hospitals consist of sections such as ICU, CCU, and specialist physicians and surgeons departments. Patients can be allocated to each of these sectors based on their physical health status and their requirements. To calculate $\lambda'_{k,j}$, the patients, who have been allocated to hospital j based on their health status and their requirements, are allocated to the low-level sections if condition $\sum_{k=1}^{K} p_{k,j} = 1$ is observed. So, Constraint 7 is as below:

$$\lambda'_{k,j} = \lambda_j w_{k,j} p_{k,j}. \tag{18}$$

To calculate $\lambda''_{z,j}$, it is assumed that the patients, who have been assigned to the low-level sections with probability $p'_{z,j}$, will need the high-level sections if condition

$\sum_{z=1}^{Z} p'_{z,j} = 1$ is observed; therefore, these patients are instructed from low-levels to the high-levels. (The idea of hierarchical structure inside hospitals has been used.)

$$\lambda''_{z,j} = \sum_{k=1}^{K} \left[\lambda'_{k,j} x'_{k,z,j} w'_{z,j} p'_{z,j}\right]. \tag{19}$$

However, if patients have serious situation, they should be allocated directly to the high-level sections without being allocated to the low-level sections if condition $\sum_{z=1}^{Z} p''_{z,j} = 1$ is observed. In this case, the patients, who have been assigned to the low-levels, will be deducted from the patients entered the hospital; the remaining patients are those who will need only high-level sections.

$$\left[\left(\lambda_j - \sum_{k=1}^{K} \lambda'_{k,j}\right) \cdot \left(w''_{z,j} p''_{z,j}\right)\right]. \tag{20}$$

Then, the equation expressed in Constraint 8 is as follows:

$$\lambda''_{z,j} = \sum_{k=1}^{K} \left[\lambda'_{k,j} x'_{k,z,j} w'_{z,j} p'_{z,j}\right] + \left[\left(\lambda_j - \sum_{k=1}^{K} \lambda'_{k,j}\right) \cdot \left(w''_{z,j} p''_{z,j}\right)\right]. \tag{21}$$

Some entered patients to the hospitals are dissuaded due to various factors, such as overcrowding in the hospitals, long queues, or other personal factors. In this case, the patients, who have been assigned to the low and high levels, will be deducted from the patients entered the hospital; the remaining patients are those who have been dissuaded:

$$N_{\text{Dissuaded}} = \left[\left(\lambda_j - \sum_{k=1}^{K} \lambda'_{k,j} - \sum_{z=1}^{Z} \lambda''_{z,j}\right)\right]. \tag{22}$$

Constraints 9 and 10 assure that the arrival rate for each low-level and high-level sections must be less than their resource capacity (because the shortage is not allowable). Constraints 11 and 12 indicate that resource capacity inside the low-level and high-level sections is bounded. Low levels and high levels inside each hospital are already predetermined and equal for all hospitals; for example, we can say that we are opening two hospitals with two low-level and two high-level sections. Because these levels are predetermined, the various sections of the hospitals will be opened when hospitals are open. Therefore, $y'_{k,j}$, $y''_{z,j}$ are the location parameters. Constraint 13 indicates that for each type of the low-level resources (such as bed,..), the sum of each resource in all of the low-level sections is equal to the overall capacity of the hospital; Constraint 14 expresses that for each type of the high-level resources (such as ICU bed,...), the sum of each resource in all of the high-level sections is equal to the overall capacity of the hospital. In addition, Eq. (15) expresses that k, z are equal in all of the hospitals, but r^L_j and r^H_j: resources in the low-level sections and those in the high-level sections are not equal in all hospitals; so, this matter exerts effects on the total setup cost. The resources inside the low-level sections of each hospital are different from the resources inside the low-level sections of the other hospital, and this is also true for high-level section resources. Therefore, this will affect the hospital startup costs (hospitals which have more resources, more consideration for the customer's convenience, and more startup costs). Since we use this criterion to calculate the setup costs of the hospitals, a criterion is obtained by gathering the total capacity of all resources in all low and high-level parts of j hospital. Through multiplying $c^{C'}_j$ by the cost factors, c^C_j will be obtained. Therefore, the setup costs of the hospitals are composed of three categories of costs:

1. The cost of land, labor, equipment, and materials used in the construction of facilities;
2. Average transportation costs of all clients to each hospital (average distance multiplied by the average demands of all customers from each hospital) are considered as a base cost of building hospitals to consider the welfare of the clients and to make the model closer to the reality;
3. A cost that is based on the available resources in each level of the hospitals.

4. Solution methods

4.1. Genetic algorithm (GA)

Genetic algorithm is an optimization technique based on natural evolution. The basic principles of GAs were introduced by Holland [29].

4.1.1. Encoding

In this problem, the permutation encoding is used. Every chromosome is a string of integer numbers representing the position of the hospitals in a sequence.

4.1.2. Initialization

In a GA approach, as soon as a chromosome is generated, it needs a fitness value for evaluating the obtained chromosomes and assigning a fitness value to them. We should consider the model's constraints. Therefore, we apply the penalty methods. Penalty methods are one of the first approaches in GA [30]. These authors used F penalty functions in their paper to penalize infeasible solutions by reducing their fitness values in proportion to the degree of their violation. The penalty method transforms a constrained problem to an unconstrained one as illustrated below. In this paper, the additive penalty function is used as shown in Eq. (23). The $v(x)$ is penalty's value, and if no violation occurs, $v(x)$ will be zero; otherwise, it will take a positive amount. Since different constraints may differ in different dimensions, it is essential to normalize all the constraints. After normalizing the constraints, violations find the same dimension and the total violation can be computed as the mean of all normalized constraint violations. For instance, a constraint like $g(x) \leq b$ can be normalized as $\frac{g(x)}{b} - 1$. With regard to the presented chromosome of the model, only constraints (5) and (6), and (9)-(12) may be violated. In such a situation, these constraints are normalized and formulated as follows:

$$F(x) = \begin{cases} f(x) & x \in \text{feasible region} \\ f(x) + v(x) & \text{otherwise} \end{cases} \tag{23}$$

4.1.3. Selection rule

From the population, the chromosomes are selected to be parents to crossover and produce offsprings. In the selection methods, tournament selection is the most popular method in genetic algorithms. Moreover, binary tournament selection is probably the most popular. In this method, to select each candidate solution, we select a random subset of solutions (k) from the original population, and then we select the best solution out of this subset (in binary tournament selection $k = 2$).

4.1.4. Crossover operator

The crossover of enumerated chromosomes depends upon their representation type. For chromosomes using repeating representation, there is no restriction on the gene value before and after crossover, because any gene may have any value from the alphabet independent of the value of all other genes. This type of crossover swaps parts of two chromosomes between each of two mates involved in the reproduction process. This type of crossover is shown in Figure 1. The enumerated chromosomes using the unique representation have a strong constraint; after the crossover operation, they must not contain genes with equal values; otherwise, if there are duplicate values in the new offspring, such an individual will be in conflict with the conditions of the problem.

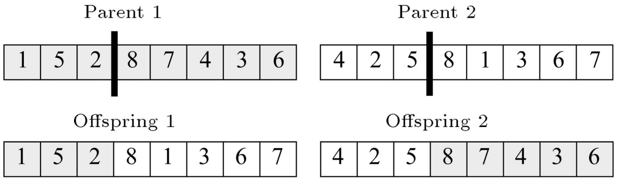

Figure 1. Enumerated chromosomes can have repeating representation, so after the crossover operation, the offspring can contain duplicate gene values.

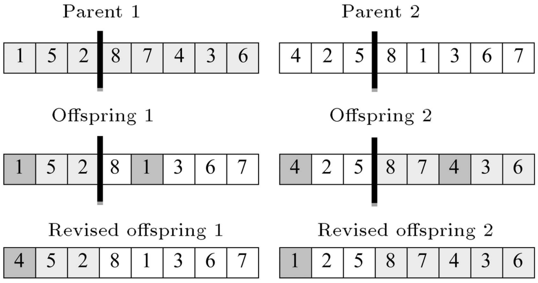

Figure 2. An enumerated chromosome has unique representation, and after crossover, the new offspring never contain duplicate gene values.

With the crossing point between the third and fourth genes, we will create two new offsprings; the new offsprings do not correspond to the main condition of the problem. The first offspring contains double 1s, and the second one contains double 4s. According to the problem conditions, each hospital must be presented only once. Therefore, we need to invent a new type of crossover for enumerated genes with unique representation. The first section of mate 1 consisting of strings 1-5-2 is moved into offspring 1. The second section of mate 2 (8-1-3-6-7) cannot be moved directly into offspring 1. Therefore, swapping takes place. In the first section of offspring 1, exchange takes place: 1 is removed and 4 is replaced. After such exchanges, there will not be duplicated genes in the new offspring. This procedure will take place for offspring 2. Figure 2 shows the procedure.

4.1.5. Mutation

Swap mutation is used in this paper; Figure 3 shows the action.

4.1.6. Updating population

The procedure of population merging in this research is that the population of the previous generation with new generated population is merged and examined by crossover and mutation operators, and the eventual better fitness population of the size of NPOP is chosen as the new generation.

4.1.7. Stopping criteria

The maximum number of iterations must be met to stop the algorithm.

4.2. Simulated Annealing algorithm (SA)

Simulated annealing is a probabilistic discrete optimizing method proposed by Krikpatrick et al. [31], and then developed by Cerny [32] for finding the global minimum of a cost function that may possess several local minimals.

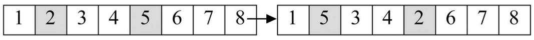

Figure 3. Swap mutation.

4.2.1. Starting temperature

Starting temperature must be hot enough to allow moves toward almost neighbourhood state (else, we are in danger of implementing hill-climbing). It must not be so hot that we conduct a random search for a period of time.

4.2.2. Initial solution

It is considered randomly and has permutation encoding.

4.2.3. Create neighbors

To create a neighborhood of the current solution, several methods are used that are common with mutation operators in genetic algorithm. Algorithm randomly selects a form among these methods. Three types of mutation operators for permutation are as below:

1. *Swap mutation:* Figure 3 shows the swap mutation,
2. *Insertion mutation:* Figure 4 shows the insertion mutation,
3. *Inversion mutation:* Figure 5 shows the inversion mutation.

4.2.4. Iterations at each temperature

Theory states that we should allow enough iterations at each temperature so that the system stabilizes at that temperature. The constant number of iterations at each temperature is used in this paper.

4.2.5. Temperature decrement

Temperature is initially high, and many moves are accepted. We decrease the temperature slowly, and accept less bad moves at each temperature level until very low temperatures at which the algorithm becomes a greedy hill-climbing algorithm. One of the decrement methods is of linear type, and the function of decrement is $T = \alpha * T$.

4.2.6. Stopping criteria

It is usual to let the temperature decrease until it reaches zero; however, this can make the algorithm run for a lot longer. Then, it is considered that the final temperature is zero and the stopping criterion is suitable number of iterations.

To make Δc without dimension, we must divide it into the cost of either initial solution or new solution. In this paper, we divide it into the initial solution cost. In this way, Δc is being independent of temperature, and algorithm can start at lower temperatures; indeed, the dependence of T0 and TF on the cost scale will become minor. Eq. (24) shows the above issue:

$$\Delta c = \frac{F(x_{\text{new}}) - F(x_{\text{old}})}{F(x_{\text{old}})}. \tag{24}$$

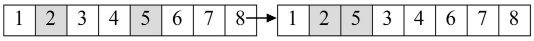

Figure 4. Insertion mutation.

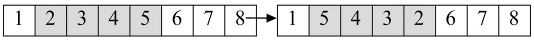

Figure 5. Inversion mutation.

4.3. Hybrid of GA and SA (HGSAA)

The meta-heuristic algorithms can be combined together, which is among one of their many advantages. This subject is important. When solving models by only one meta-heuristic algorithm, we will realize that our algorithm consists of some weaknesses, and we will consider some techniques to solve the problem [33]. In this section, we have tried to get help from two meta-heuristic algorithms consisting of genetic and simulated annealing algorithms; we have proposed a hybrid algorithm to solve our model. Hybrid algorithm used in this paper applies a genetic algorithm to find the initial solution, and its final solution is the initial solution of simulated annealing. As we know, the genetic algorithm is not sensitive to the initial solution, and the final result is not particularly impressed by the initial solution. However, we know that if the SA starts with a good solution due to its high response, it will surely converge to a good solution. In this way, it will be closer to favorable solutions. The convergence of the hybrid algorithm used in this section is proved, its reason is clear, and it consists of two popular algorithms. The genetic algorithm explores a range of solution space to find the solution; consequently, it investigates the points of the feasible solution region of the problem that gets less attention. This increases the probability of finding a suitable initial solution. Then, the hybrid algorithm offers the solution to the SA and makes new neighbors for the initial solution and explores the surrounding areas. More significantly, it will move further towards the optimum solutions.

4.3.1. Structure of the proposed hybrid algorithm (HGSAA)

In this section the structure of the hybrid algorithm is presented.

1. Begin,
2. Determine *Parameters*,
3. Generate *Initial solution by genetic algorithm*,
4. Repeat *for Maxit=1: MaxIT*,
5. Generate *new neighborhood solution*,
   By simulated annealing algorithm,
6. Check *Stop condition*,
7. Show *Best answer*,
8. End.

4.4. Parameter tuning

The parameters of GA are the number of population (NPOP), the Probability of Crossover (PC), the Probability of Mutation (PM), and the number of generation (MaxIT); the parameters of SA are the number of generation (MaxIT), the number of iterations at each temperature (MaxIPT), initial temperature (T0), and temperature reduction rate (α). These parameters can take values in different ranges and the optimum result depends on the combination values of these parameters. Some of the combination values of these parameters are given in Tables 1 and 2.

In order to obtain a better solution in shorter time, a parameter adjustment approach is applied in this section to tune the control parameters. While classical statistical optimization tools are alternatives to tune the control parameters, the orthogonal arrays of the Taguchi method [34] are usually performed to study more decision variables in smaller number of experiments. Taguchi L_9 orthogonal array is used for four factors of each algorithm, each at three levels with a total of nine observations on the response. Five examples of different sizes were generated and used five times for different nine combinations of each of the factor levels, where the stopping criterion is met in each algorithm. By parameter tuning of the GA and SA, the parameters of HGSAA are tuned automatically, and there is no need for parameter tuning of hybrid algorithm. Figures 6 and 7 show the mean S/N ratio plots for each factor level in problems, respectively. Since the goal is to maximize S/N ratio, the best combination of the factor levels for different problem sizes can be easily obtained based on these figures. Tables 3 and 4 contain all of the best factor level combinations for all of the problems.

Table 1. The factor levels for GA algorithm.

Parameters	Low	Medium	High
MaxIT	50	100	150
NPOP	150	200	250
PC	0.6	0.7	0.8
PM	0.1	0.3	0.5

Table 2. The factor levels for SA algorithm.

Parameter	Low	Medium	High
MaxIT1	50	200	250
MaxIPT	30	40	50
T0	800	1000	1250
α	0.8	0.9	0.99

Table 3. The best factor level combination for GA algorithm.

Parameters	NPOP	MaxIT	PC	PM
Parameter values	250	150	0.8	0.3
Best levels	3	3	3	2

Table 4. The best factor level combination for SA algorithm.

Parameters	MaxIT	MaxIPT	T0	α
Parameter values	150	40	1000	0.99
Best levels	1	2	2	3

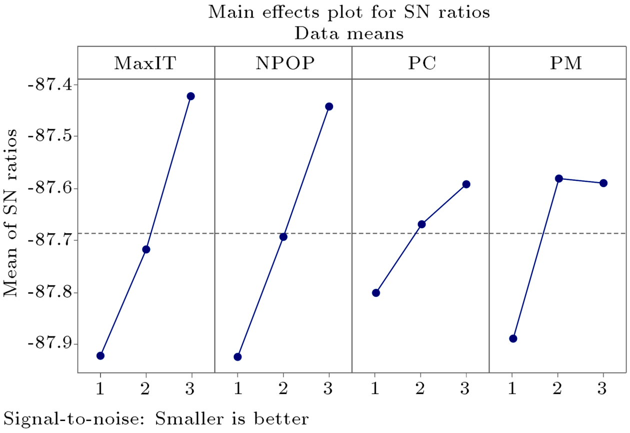

Figure 6. The mean S/N ratio plot for the factors of GA algorithm.

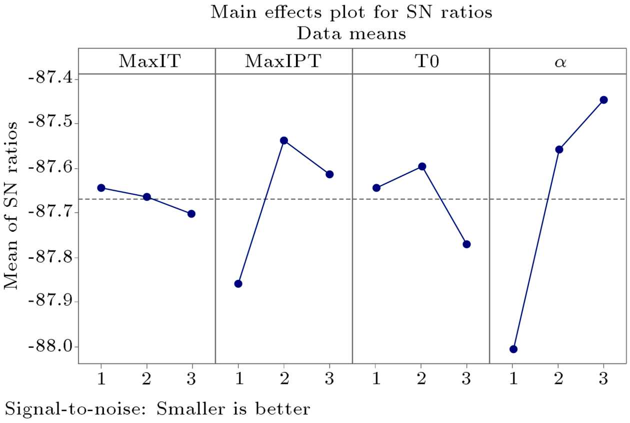

Figure 7. The mean S/N ratio plot for the factors of SA algorithm.

5. Validation of the Model

For validation of the model, first, 10 test problems in small and average sizes were generated, because the GAMS Software is incapable of solving large-sized problems and needs long-time runs for large-sized problems. Then, these problems were solved via optimization software GAMS and meta-heuristic algorithms. The meta-heuristic algorithms are coded and compiled in Matlab (R2013a), and the computational experiment is performed on a PC with 32-bit operating system (Windows 7) and with Intel, Core i7 CPU, 1.6 GHz Processor and 4GB of RAM. Every example was run 3 times in each meta-heuristic algorithm, and their average was compared with the results of GAMS software in Table 5. In this section, the performances

Table 5. Comparison of the Objective Function Values (OFV) of GA and SA algorithms and GAMS software.

Order	M	N	OFV (GA)	OFV (SA)	OFV (GAMS)	Site of opened facilities	Number of opened facilities
1	3	2	6454	6454	6454	[1 2]	2
2	5	2	3605	3605	3605	[1 2]	2
3	7	3	17960	17969	17960	[3 2]	2
4	10	3	21088	21102	21108	[2 3]	2
5	12	3	27053	27053	27047	[1]	1
6	15	4	24628	24650	24670	[3 4]	2
7	17	4	20706	20744	20718	[4 1 3]	3
8	20	5	27675	27766	27711	[1 5 3]	3
9	22	5	24599	24648	24652	[3 2]	2
10	25	6	29590	29656	30578	[4 5 1]	3

of the parameters of tuned GA and SA and GAMS are compared using a statistical test (ANOVA).

5.1. Statistical comparison

Analysis of variance (ANOVA) is used to determine whether there are any significant differences between the means of two or more independent groups. The ANOVA tests are the null hypothesis whose samples in two or more groups are drawn from populations with the same mean values.

ANOVA test only tells us that at least two groups are different. Determining which of these groups differ from each other is important, and we can do this using post-hoc tests. One-way ANOVA consists of 3 parts: First, testing the variance equality of groups if the test does not reject the equality of variances; in the second part, the results of the tests on the means of the equality of group are considered in the case of variances' equality. If the variance equality is rejected, testing the equality of means is done in the case of inequality of variances; if the hypothesis of equality of means is not rejected, the analysis will be completed, and indicates that there is no difference between the groups. But, if the assumption of the equality of means between treatments is rejected, the need of looking for differences is indicated. In the third part, to determine which treatments are significantly different, post-hoc tests will be used. Some of these tests, including Dunnett, Duncan, Tuckey, and LSD, are used in the case of equality of variances; Dunnett's C, Games-Howell, Dunnent T3, Tmahanes's T2 are used in the case of the inequality of variances. So, ten test problems are generated and solved by GA, SA, and GAMS software products. In this regard, one-way ANOVA is performed at 95% significant level for the comparison of each algorithm's results and the results of GAMS software, while for the mean of fitness value comparison, the hypotheses are as follows:

H₀: $\mu_{\mathrm{GA}} = \mu_{\mathrm{SA}} = \mu_{\mathrm{GAMS}}$

H₁: Otherwise

Table 6. Test of the homogeneity of variances for the OFVs of GA, SA and GAMS software.

Levene statistic	*df*1	*df*2	Sig.
0.000	2	27	1.000

A comparison between the results of the solved sample problems was done. The results of tests for the equality of variances and means for fitness values are presented in Tables 6 and 7.

As the results show, the significance of the equality of variances is greater than 0.05; therefore, the assumption of the equality of variances will be accepted. Also, the significance for the equality of means is greater than 0.05; so, the assumption of the equality of means is also accepted. Therefore, it can be concluded that the proposed algorithms in the significance level of 0.95 are similar to the results of GAMS software.

6. Numerical experiment and results

In this section, several test problems with different sizes (small, medium, and large) are solved to evaluate the performance of the three presented meta-heuristic algorithms. In total, 10 test problems were generated randomly and these problems were run for three times. Therefore, thirty runs were done with each algorithm, and the results are shown in Table 8.

To evaluate the performance of the mentioned parameter tuned algorithms, one-way ANOVA at 95% significant level is performed, while for the mean fitness value comparison, the hypotheses are as below:

H₀: $\mu_{\mathrm{GA}} = \mu_{\mathrm{SA}} = \mu_{\mathrm{GSHAA}}$,

H₁: One of them is not equal,

Table 7. ANOVA table for the equality of means for the OFVs of GA, SA and GAMS software.

	Sum of squares	***df***	**Mean square**	***F***	**Sig.**
Between groups	72877.400	2	36438.700	0.000	1
Within groups	2.123E9	27	78641054.244		
Total	2.123E9	29			

Table 8. Comparison of the OFVs and CPU times of the proposed algorithms.

Run order	M	N	**GA**		**SA**		**HGSAA**	
			OFV	**CPU**	**OFV**	**CPU**	**OFV**	**CPU**
1	5	2	3605	69.78	3605	13	3605	95
2			3605	80.2	3605	16	3605	87
3			3605	73	3605	13.87	3605	90
4	10	3	16047	86	16052	25	16045	118
5			16047	85	16053	17	16051	118
6			16052	84	16045	16	16045	116
7	15	4	24633	111	24640	21	24635	155
8			24638	120	24628	23	24628	155
9			24635	114	24655	19	24633	152
10	20	5	27691	129	27725	25	27686	170
11			27698	124	27766	22	27692	167
12			27690	127	27770	24	27685	172
13	25	6	29590	203	29656	32	29594	227
14			29609	164	29608	27	29584	224
15			29588	169	29668	36	29590	228
16	30	5	33850	149	33919	33	33797	247
17			33817	148	34004	29	33843	235
18			33853	149	34010	28	33804	264
19	35	7	29214	318	29432	85	29220	325
20			29271	345	29408	85	29232	339
21			29232	295	29469	86	29193	369
22	40	7	52176	210	52764	35	52121	354
23			52198	213	52981	92	52098	382
24			52063	202	52713	91	52144	437
25	50	10	47360	328	47530	53	47305	382
26			47333	323	47855	51	47316	485
27			47265	341	48016	51	47325	494
28	60	15	47660	523	49606	97	47734	789
29			47790	525	49494	88	47760	656
30			47634	560	49714	85	47714	650

and for the mean run time comparison, the hypotheses are:

$\mathbf{H_0}$: $\mu_{\text{GA}} = \mu_{\text{SA}} = \mu_{\text{GSHAA}}$,

$\mathbf{H_1}$: One of them is not equal.

A comparison between the results of solving sample problems was done. The results of tests for the equality of variances and means for fitness values are presented in Tables 9 and 10.

Table 9. Test of the homogeneity of variances for the OFVs of GA and SA.

Levene statistic	***df*1**	***df*2**	**Sig.**
0.017	2	87	0.984

As the results show, the significance of the equality of fitness function variances is greater than 0.05; therefore, the assumption of the equality of variances will be accepted. Also, the significance of the equality of fitness function means is greater than 0.05; therefore, the assumption of the equality of means is also accepted. Therefore, it can be concluded that the proposed algorithms in the significance level of 0.95

Table 10. ANOVA table for the equality of means for the OFVs of GA and SA.

	Sum of squares	*df*	Mean square	*F*	Sig.
Between groups	2573993.756	2	1286996.878	0.006	00.994
Within groups	18594310412.067	87	213727705.886		
Total	18596884405.822	89			

Table 11. Test of the homogeneity of variances for the CPU times of GA and SA.

Levene statistic	*df* 1	*df* 2	Sig.
18.514	2	87	0.000

are similar, and there are not significant differences between the fitness means. The comparison of the fitness' mean figure is presented in Figure 8. The results of tests for the equality of variances and means of run times are presented in Tables 11 and 12.

As the results show, the significance of the equality of variances is smaller than 0.05; therefore, the assumption of the equality of variances will be rejected. Moreover, a test for the equality of means in the case of inequality of variances is done. The significance of the equality of CPU means is smaller than 0.05; therefore, the assumption of the equality of means is rejected. The assumption of the equality of means between treatments is rejected, indicating the need of looking for differences. To determine which treatments are significantly different, a post-hoc test in the case of inequality of variances will be done. These tests are shown in Table 13.

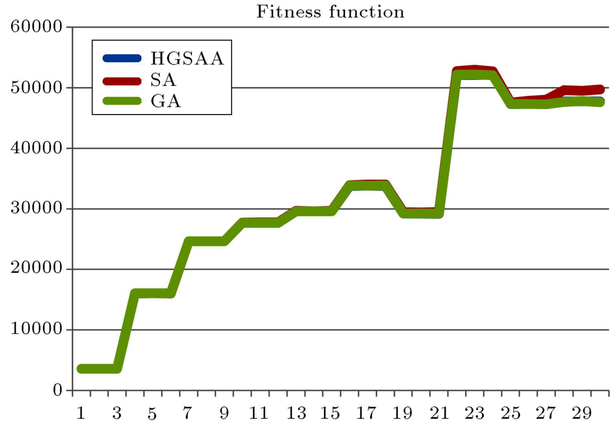

Figure 8. Comparison of the OFVs of the proposed algorithms .

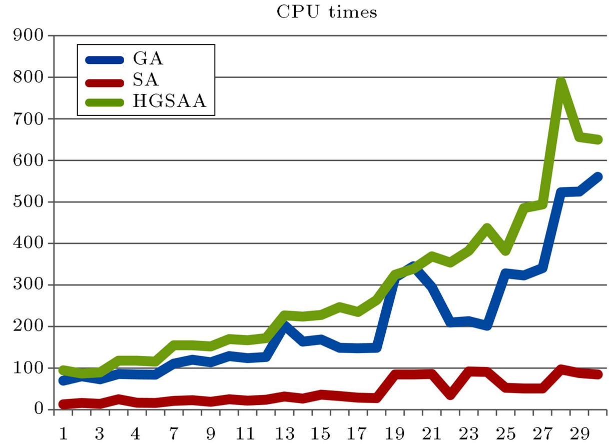

Figure 9. Comparison of the CPU times of the proposed algorithms.

Therefore, it can be concluded that the proposed algorithm's run times at the significance level of 0.95 are not similar, and there are significant differences between the mean run times in SA and GA, SA and HGSAA. But, there is no significant difference between GA and HGSAA run times. The comparison of means figure is presented in Figure 9. There is no significant difference in objective function means.

7. Conclusions and future study

In this research, a mixed integer programming model was developed for a capacitated location-allocation problem in the health care systems. Most real-world engineering problems consist of different objectives simultaneously. Therefore, a trade-off among these objectives is very important to us. We tried to balance our objective functions, including facility setup, operation costs, and patient costs by considering weights for two objective functions; so, we will allow the decision-

Table 12. ANOVA table for the equality of means for the CPU times of GA and SA.

	Sum of squares	*df*	Mean square	*F*	Sig.
Between groups	945154.264	2	472577.132	26.344	0.000
Within groups	1560653.620	87	17938.547		
Total	2505807.884	89			

Table 13. Multiple comparisons for the CPU times of the proposed algorithms.

	(*I*) algorithms	(*J*) algorithms	Mean difference (*I*-*J*)	Std. error	Sig.	95% confidence interval	
						Lower bound	Upper bound
Games-Howell	GA	SA	168.30367*	26.05526	0.000	104.2309	232.3764
		HGSAA	-77.13400	42.01618	0.168	-178.3796	24.1116
	SA	GA	-168.30367*	26.05526	0.000	-232.3764	-104.2309
		HGSAA	-245.43767*	33.81529	0.000	-328.7339	-162.1414
	HGSAA	GA	77.13400	42.01618	0.168	-24.1116	178.3796
		SA	245.43767*	33.81529	0.000	162.1414	328.7339
Dunnett C	GA	SA	168.30367*	26.05526		103.9564	232.6510
		HGSAA	-77.13400	42.01618		-180.8992	26.6312
	SA	GA	-168.30367*	26.05526		-232.6510	-103.9564
		HGSAA	-245.43767*	33.81529		-328.9495	-161.9258
	HGSAA	GA	77.13400	42.01618		-26.6312	180.8992
		SA	245.43767*	33.81529		161.9258	328.9495

* The mean difference is significant at the 0.05 level.

makers to decide based on the patient satisfaction or operation costs of the hospitals. Our main purpose is to establish a mathematical model that is more practical in the real world. Opening costs of the hospitals are based on three categories: the patient costs, the fixed costs of establishing health centers, and finally the costs based on the available resources in each level of hospitals, respectively. The model aims to open an appropriate number of hospitals and determine the best sites for them to satisfy the patients by allocating them to the closest hospitals and simultaneously minimizing the operating costs of hospitals. There are two levels of service inside the hospitals: low-level and high-level services, and each level has different sections to visit various types of patients and the sections at different levels have different resources; therefore, the other goal of the models is to determine the resource capacity at different sections, and consequently system capacity of opened facilities. Two meta-heuristic algorithms (GA and SA) and their combination (HGSAA) have been executed for the produced test problems. One-way ANOVA was employed where the performances were compared in terms of fitness function values and CPU run times for the comparison purpose. The results of several numerical examples in small-, medium-, and large-sized categories showed that there is no significant difference in the objective function means. Moreover, there is a significant difference in the run time means of the mentioned three algorithms, and the parameters of tuned GA and HGSAA are a better procedure than the parameters of tuned SA in terms of average fitness value. However, the SA has a better result in terms of average run time compared to that of the two other algorithms.

The following approaches can be proposed for the future research studies:

1. Considering another objective function (for example, maximal covering location problem) instead of P-median objective function;
2. Considering the multi-objective function problem and solving it by appropriate meta-heuristic algorithms (such as NSGA-II and NRGA);
3. Considering some of the parameters, fuzzy or random, and making the model closer to the reality; shortages are involved in systems;
4. Employing other meta-heuristic algorithms or heuristic algorithms or hybrid of heuristic and meta-heuristic algorithms to solve the model and investigate their efficiencies;
5. Developing other queuing system rather than M/M/C/K;
6. Developing heuristic approach instead of generating random data in the initial segment.

References

1. Current, J., Daskin, M. and Schilling, D. "Discrete network Location models", In *Facility Location: Applications and Theory*, Z. Drezner, H.W. Hamacher, Eds., Springer, Heidelberg, **14**, pp. 80-118 (2002).
2. Balinski, M.L. "Integer programming, methods, uses, computations", *Manag Sci.*, **12**, pp. 253-313 (1965).
3. Weber, A., *Theory of the Location of Industries*, Über den Standort der Industrien, TÜbingen, University of Chicago Press (1929).

4. Drezner, E., *Facility Location, A Survey of Applications and Methods*, Springer (1995).

5. Dantrakul, S., Likasiri, C. and Pongvuthithum, R. "Applied p-median and p-center algorithms for facility location problems", *Expert Syst. Appl.*, **41**(8), pp. 3596-3604 (2014).

6. An, Y., Zeng, B., Zhang, Y. and Zhao, L. "Reliable *p*-median facility location problem: two-stage robust models and algorithms", *Transport Res. B-Meth.*, pp. 54-72 (2012).

7. Toregas, C., Swain, R., Revelle, C. and Bergman, L. "The location of emergency service facilities", *Oper Res.*, **19**, pp. 1363-1373 (1971).

8. Teresinha Arns Steiner, M., Datta, D., Steiner Neto, P.J., TadeuScarpin, C. and Figueira, J.R. "Multi-objective optimization in partitioning the healthcare system of Parana State in Brazil", *Omega*, **52**, pp. 53-64 (2015).

9. Tate, W.L., Ellram, L.M., Schoenherr, T. and Petersen, K.J. "Global competitive conditions driving the manufacturing location decision", *Bus Horiz.*, **57**(3), pp. 381-390 (2014).

10. Tohidi, H. "Mathematical modeling of optimal multi fuzzy locations of facilities based on the assumed step distance among them in a convex set", *Appl. Math. Model.*, **39**(23-24), pp. 7442-7451 (2015).

11. Sainathuni, B., Parikh, P.J., Zhang, X. and Kong, N. "The warehouse-inventory-transportation problem for supply chains", *Eur. J. Oper. Res.*, **237**(2), pp. 690-700 (2014).

12. Hakimi, S. "Optimum location of switching centers and absolute centers and median of a graph", *Oper. Res.*, **12**, pp. 450-59 (1964).

13. Guerriero, F., Miglionico, G. and Olivito, F. "Location and reorganization problems: The Calabrian health care system case", *Eur. J. Oper. Res.*, **250**(3), pp. 939-954 (2016).

14. Ghaderi, A. and Jabalameli, M.S. "Modeling the budget-constrained dynamic uncapacitated facility location-network design problem and solving it via two efficient heuristics: A case study of health care", *Math. Comput Model.*, **57**(3-4), pp. 382-400 (2013).

15. Vatsa, A.K. and Jayaswal, S. "A new formulation and benders decomposition for the multi-period maximal covering facility location problem with server uncertainty", *Eur. J. Oper. Res.*, **251**(2), pp. 404-418 (2016).

16. Baray, J. and Cliquet, G. "Optimizing locations through a maximum covering/p-median hierarchical model: Maternity hospitals in France", *J. Bus Res.*, **66**(1), pp. 127-132 (2013).

17. Mohammadi, M., Dehbari, S. and Vahdani, B. "Design of a bi-objective reliable healthcare network with finite capacity queue under service covering uncertainty", *Transport Res. E-LOG.*, **72**, pp. 15-41 (2014).

18. Cooper, L. "Location-allocation problems", *Oper. Res.*, **11**, pp. 331-344 (1963).

19. Damgacioglu, H., Dinler, D., Ozdemirel, N.E. and Iyigun, C. "A genetic algorithm for the uncapacitated single allocation planar hub location problem", *Comput. Oper Res.*, **62**, pp. 224-236 (2015).

20. Kratica, J., Dugošija, D. and Savić, A. "A new mixed integer linear programming model for the multi level uncapacitated facility location problem", *Appl Math Model*, **38**(7-8), pp. 2118-2129 (2014).

21. Alizadeh, M., Mahdavi, I., Mahdavi-Amiri, N. and Shiripour, S. "A capacitated location-allocation problem with stochastic demands using sub-sources: An empirical study", *Appl Soft Comput.*, **34**, pp. 551-571 (2015).

22. Zhou, J. and Liu, B. "Modeling capacitated location-allocation problem with fuzzy demands", *Comput. Ind Eng.*, **53**, pp. 454-468 (2007).

23. Marinakis, Y. "An improved particle swarm optimization algorithm for the capacitated location routing problem and for the location routing problem with stochastic demands", *Appl Soft Comput.*, **37**, pp. 680-701 (2015).

24. Sahin, G. and Sural, H. "A review of hierarchical facility location models", *Comput. Oper. Res.*, **34**, pp. 2310-2331 (2007).

25. Zanjirani Farahani, Z., Hekmatfar, M., Fahimnia, B. and Kazemzadeh, N. "Hierarchical facility location problem: Models, classifications, techniques, and applications", *Comput. Ind Eng.*, **68**, pp. 104-117 (2014).

26. Shavandi, H., Mahlooji, H., Eshghi, K. and Khanmohammadi, S. "A fuzzy coherent hierarchical location-allocation model for congested systems", *Sci. Iran.*, **13**(1), pp. 14-24 (2006).

27. Shavandi, H. and Mahlooji, H. "Fuzzy hierarchical location-allocation models for congested systems", *Int. J. Ind. Syst. Eng.*, **1**(2), pp. 171-189 (2007).

28. Alinaghian, M., Hejazi, S.R. and Bajoul, N. "A novel hierarchical model to locate health care facilities with fuzzy demand solved by harmony search algorithm", *IJSOM.*, **1**(2), pp. 245-259 (2014).

29. Holland, J., *Adaption in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, The University of Michigan Press (1975).

30. Yeniay, Ö. and Ankare, B. "Penalty function methods for constrained optimization with genetic algorithms", *Math. Comput Appl.*, **10**(1), pp. 45-56 (2005).

31. Krikpatrick, S., Gelatt, C.D. and Vecchi, M.P. "Optimization by simulated annealing", *Science*, **220**, pp. 671-680 (1983).

32. Cerny, V. "A thermo dynamical approach to the travelling salesman problem: An efficient simulation algorithm", *J. Optim Theory Appl.*, **45**, pp. 41-51 (1985).

33. Gen, M. and Cheng, R., *Genetic Optimization and Engineering Optimization*, Wily Interscience by Wily John and Sons, New York (2000).

34. Ross, R.J., *Taguchi Techniques for Quality Engineering*, McGraw-Hill, New York (1989).

Biographies

Mahsa Pouraliakbarimamaghani was born in 1989. She is an MS student in Industrial Engineering at the Kharazmi University, Tehran, Iran. She is a member of brilliant students society in Iran. Her research interest areas are health care systems, location problems, supply chain, mass casualty events management, and meta-heuristics. She received her BS degree in Industrial Engineering (Industrial Production) from Urmia University of Technology, Urmia, Iran, in 2012.

Mohammad Mohammadi is an Associate Professor in the Department of Industrial Engineering at the Kharazmi University, Tehran, Iran. He received his BS degree in Industrial Engineering from Iran University of Science and Technology, Tehran, Iran, in 2000, and his MS and PhD degrees in Industrial Engineering from Amirkabir University of Technology, Tehran, Iran, in 2002 and 2009, respectively. His research and teaching interests include applied operations research, sequencing and scheduling, production planning, time series, meta-heuristics, and supply chains.

Abolfazl Mirzazadeh is an Associate Professor of Industrial Engineering at the Kharazmi University in Iran. He earned his PhD degree from Amirkabir University of Technology in Iran. His interest research areas are uncertain decision-making, production/inventory control, supply and operations management, development of quality management and problem-solving tools, and stochastic processes. He has more than 70 research papers published or forthcoming in journals. He earned the second rank on the inflationary inventory research studies on Scopus report in 2014. He is now the editor-in-chief of the International Journal of Supply and Operations Management.