International Journal of Industrial Engineering & Production Research (2017)

December 2017, Volume 28, Number 4
pp. 403- 427

DOI: 10.22068/ijiepr.28.4.403 http://IJIEPR.iust.ac.ir/

Multi-Objective Evolutionary Algorithms for a Preventive Healthcare Facility Network Design

Keyvan Roshan, Mehdi Seifbarghy & Davar Pishva*

Keyvan Roshan, *School of Industrial Engineering, Islamic Azad University, South Tehran Branch*
Mehdi Seifbarghy, *Department of Industrial Engineering, Alzahra University*
Davar Pishva, *Faculty of Asia Pacific Studies, Ritsumeikan Asia Pacific University, Beppu, Japan*

KEYWORDS

Multi-objective Preventive healthcare problems (MOPHPs),
Queuing system,
Multi-objective simulated annealing (MOSA),
NSGA-II,
NRGA,
Taguchi method.

ABSTRACT

Preventive healthcare aims at reducing the likelihood and severity of potentially life-threatening illnesses by means of protection and early detection. In this paper, a bi-objective mathematical model is proposed to design a network of preventive healthcare facilities in which each facility acts as M/M/1 queuing system so as to minimize total travel and waiting time as well as establishment and staffing cost. The number of facilities to be established, the location of each facility, and the level of technology for each prospect facility are provided as the main determinants of a healthcare facility network. Since the developed model of the problem is of an NP-hard type, tri-meta-heuristic algorithms are proposed to solve the problem. Initially, Pareto-based meta-heuristic algorithm, which is called multi-objective simulated annealing (MOSA), is proposed to solve the problem. Subsequently, obtained results are validated by means of two popular algorithms, namely non-dominated sorting genetic algorithm (NSGA-II) and non-dominated ranking genetic algorithm (NRGA). Considering that solution-quality of all meta-heuristic algorithms heavily depends on their parameters, Taguchi method is used to fine tune parameters of the employed algorithms. The computational results, obtained by implementing the algorithms on several problems of different sizes, demonstrate the reliability of the proposed methodology. It efficiently minimizes establishment and staffing costs, as well as travel and waiting time for the service, something which is directly related to the ultimate goal of managerial strategies for maximum preventive healthcare participation achievement.

1. Introduction1

Preventive health care is of utmost importance to governments, since they can make massive savings on health care expenditure and promote well-being of the society. Preventive care includes many services such as cancer screenings, vaccinations, hepatitis screenings, and smoking cessation programs. Despite the benefits of these services, their uptake is not satisfactory in many countries in the world. This

* Corresponding author: *Davar Pishva*
Email: dpishva@apu.ac.jp
Received 17 May 2017; revised 3 October 2017; accepted 15 October 2017

can be attributed to financial barriers, social issues, and some other factors. One of the most important barriers for preventive care is accessibility to proper services, which is a function of various qualitative and quantitative factors such as distance to travel, waiting time, presence of other attractive facilities (e.g., shopping malls) in the vicinity, and even cleanliness of the facilities. Statistics show that even a small improvement in people's participation can save massive amounts of money for any government and improve the well-being of the people in a society [1].

Preventive healthcare services aim at reducing the likelihood and severity of life-threatening illnesses through early detection and prevention. Effectiveness of these programs depends on mass participation level and accessibility of the facilities that provide such services to potential users. In order for such services to be effective, the preventive healthcare facilities should be easily accessible. Factors that influence accessibility include the number, type, and location of the facilities as well as assignment of the clients to these facilities. The level of participation in preventive healthcare programs is a critical determinant in terms of their effectiveness and efficiency. Preventive health care programs can save lives and contribute to a better quality of life by diagnosing serious medical conditions early. However, unlike sick people who need urgent medical attention, undergoing through preventive healthcare is not urgent. Hence, their clientele have more flexibility regarding when and where to receive such services. In order to maximize total participation in a preventive care program, it is important to incorporate how potential clients choose the facilities to patronize [2]. Effective preventive healthcare services have a significant role in reducing fatality and medical expenses in all human societies, and their level of accessibility to customers can be considered as a measure of their efficiency and effectiveness [3].

Healthcare infrastructure is essential for effective operations of healthcare systems. An efficient facility location can save cost and improve the facility utilization. It is important to update the knowledge of methods and applications to locate healthcare facilities for different purposes [4]. The Preventive Health Care Facility Location (PHCFL) problem is to identify optimal locations for preventive health care facilities so as to maximize participation. When identifying locations for preventive health care facilities, we need to consider the characteristics of the preventive health care services. First, people should have more flexibility in selecting service locations. Second, each preventive health care facility needs to have a minimum number of clients in order to retain accreditation [5]. Reliability of modeling in facility location problems is one of the most effective ways to hedge against failures of system from time to time. In reality, the combined facility location network design problem with respect to the reliability of system has several applications in industries and services such as locating health care service centers, locating gas compressor stations, and designing water-tubing networks [6]. In this paper, total time spent in receiving preventive health care services is used as an indicator for accessibility of healthcare facilities. This time includes the time it takes to go to the facility as well as the time spent at the facility while waiting and receiving the services. Fig. 1 shows a descriptive visual representation of the envisioned system, input of which serves as optimized main decision variables and the output as the desired objective.

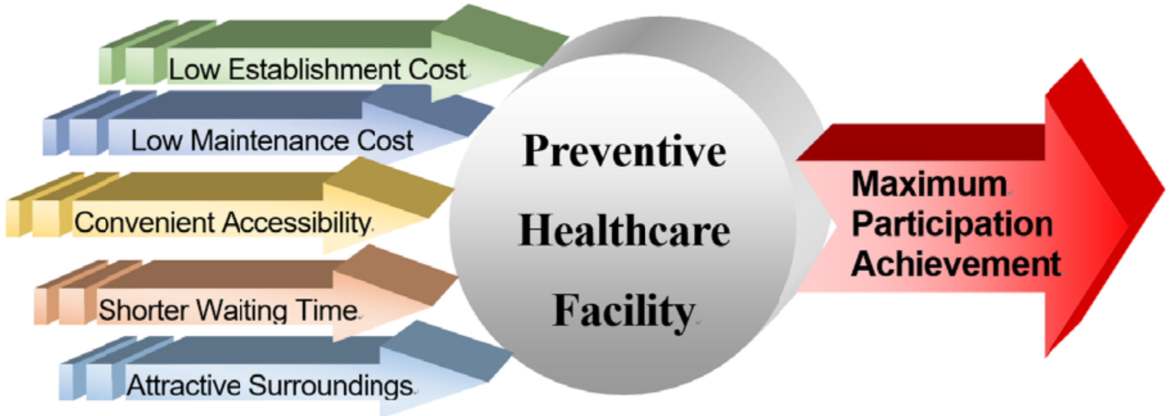

Fig. 1. Visual Representation of the Envisioned System

The number of facilities to be established, the location and the level of technology of each facility are the main determinants of the configuration of the healthcare facility network. Such an approach clearly shows the need for the development of an analytical framework for making structural decisions with regard to preventive healthcare facility networks and addressed in this paper. The methodology presented in this paper incorporates the differentiating features of preventive healthcare. A multi-objective simulated annealing (MOSA) algorithm is presented to solve the proposed model and its outcome is compared with performance of non-dominated sorting genetic algorithm (NSGA-II) and non-dominated ranking genetic algorithm (NRGA).

The rest of the paper is organized as follows. The next section provides literature reviews and cites many relevant recent research results. Section 3 describes the problem in detail and formulates it as a nonlinear programming model. Section 4 contains the development of the three meta-heuristic algorithms. The application of the proposed methodology and the statistical comparisons of the solution algorithms are studied next in Section 5 using several test problems of different sizes. Finally, conclusions are drawn and some possible future research works are recommended in Section 6.

2. Literature Review

The facility location problems (FLPs) deal with optimal location of new facilities along with their demand nodes allocations, an area for which many models have been developed under different scenarios. The term “location” relates to the modeling, formulation, and solving methodology of a class of problems that can be best described as locating facilities in some given space. The FLPs have various applications in realistic problems. One of the important scopes in this approach is facility location of preventive healthcare networks design. In practice, preventive healthcare problems (PHPs) have been used in generating substantial savings in the costs of diagnosis and therapy along with the lower capital investment [7]. Preventive programs can save lives and contribute to a better life quality by reducing the needs for radical treatments such as surgery or chemotherapy. For example, mammograms taken on a regular basis have the potential to reduce deaths from breast cancer for women between the ages of 50 and 69 by up to 40% [8]. Gornick et al. found out that 36% of breast cancer patients without a mammogram received the diagnosis of late stage cancer, whereas this ratio was 20% for the patient group who had undergone through mammography tests [9]. Preventive healthcare programs can be categorized into three groups with regard to their objectives: (i) primary prevention aims at reducing the likelihood of diseases in people with no symptoms, e.g., immunizations of healthy children; (ii) secondary prevention aims at identifying and treating people who have risk factors or are at very early stage of diseases, e.g., pap smears to detect early forms of cervical cancer; (iii) tertiary prevention aims at treating symptomatic patients in an effort to decrease complications or severity of disease, e.g., sugar control in a diabetic in order to mitigate vision and nerve problems. Flu shots, blood tests, mammograms and anti-smoking advice are among the most well-known preventive services. According to World Health Organization, although many diseases can be prevented, the current healthcare systems do not make the best use of their available resources to support preventive programs [10]. Most of these systems are based on responding to acute problems, urgent needs of patients, and pressing concerns. Preventive healthcare is inherently different from healthcare for acute problems, and in this regard, current healthcare systems worldwide fall remarkably short. For instance, only 5% of the $1.4 trillion spent on direct health care in the United States goes to preventive health measures and the promotion of general health [11]. An effective way to improve the efficiency of a regional healthcare system with limited recourses is to increase the number of people receiving preventive services, which has been an integral part of many healthcare reform programs during the past two decades [12]. Unlike the sick who need urgent medical attention, people who seek preventive services have more flexibility as to when and where to receive preventive healthcare services. Hence, accessibility of such facilities becomes an important factor for the success of a preventive healthcare program. Institute of Medicine defined access as the timely use of personal health services to achieve the best possible health outcomes [13]. According to the Institute, three groups of factors influence the individuals' use of services in healthcare: structural, financial, and personal barriers. In this paper, we focus on structural barriers that are directly related to the number, type, concentration, level of technology and location of healthcare facilities, as well as available means of transportation to the centers. This paper presents a methodology for designing a network of preventive healthcare facilities so as to minimize their establishment and staffing costs, as well as the average gross time required for the service. Such objectives affect participation in preventive healthcare programs, and there are empirical evidences suggesting that the convenience of access plays a key role in the participation. For instance, Zimmerman found through a survey that the convenience of access to the facility was a very important factor in a client's decision to have prostate cancer screening [14]. Furthermore, the main reasons given for nonattendance to mammography screening were equally divided between practical difficulties and negative attitudes towards the process [15]. A survey by Facione revealed that the perceptions of lack of access to services were related to the decrease of mammography participation [16]. Firstly, the policy maker does not control the number of people who seek the services at a given facility, i.e., the selection is totally left to the clients. Unless the services are offered at convenient locations, people will not likely participate. That is, the demand for preventive programs at population zones decreases with respect to the increase in the time that needs to be spent for receiving the associated services. In the event that people have to wait for a long time to receive the services due to limited capacity, their willingness to participate in preventive programs could decrease significantly. Therefore, the level of congestion at the facilities is a crucial factor that is incorporated into our model. The second significant factor is the apparent link between volume and quality of preventive healthcare services. For example, US Food and Drug Administration required a radiologist to interpret at least 960 mammograms and a radiology technician to perform at least 200 mammograms in 24 months to retain their accreditations [17]. Although the design of healthcare facilities has been studied for a long time, the distinguishing features of preventive healthcare are not incorporated in the prevailing models. The review by Daskin and Dean on the location of healthcare facilities, for example, makes no reference to preventive care [18]. Similarly, more general literature reviews by Berman and Krass and Marianov and Serra, which focused on public facility location problems with stochastic demand and congestion in the context of fixed versus mobile servers, do not cite any articles on preventive healthcare [19-20]. To the best of our knowledge, an article by Verter and Lapierre is the only paper that tackles the problem studied in this paper [21]. Although there are many design issues for preventive healthcare programs, our paper focuses on the configuration of a network of preventive healthcare facilities so as to minimize their establishment and staffing costs and the average total time. In representing demand elasticity, the accessibility of a facility can be modeled in terms of its proximity to the potential clients [21], the total time required for receiving the service [22], or an overall utility [23]. In either case, the shape of the utilized demand decay function represents the extent of demand elasticity. The most common demand decay functions in the literature are: the linear function [21, 23]; the exponential function [24-27]; the step function [28-29]. The most common way of incorporating congestion in facility design models is to represent the facility as a queue (e.g., M/M/1 or M/G/1) and include a capacity constraint on the level of congestion. The empirical evidence suggests that the time spent waiting (or, the level of congestion) is also a significant factor in a client's facility choice, especially in preventive healthcare [30]. Thus, to improve model realism, it is necessary to incorporate congestion into a client's decision-making process. A good example of this is the early work of Parker and Srinivasan who considered waiting time as one of the attributes in a client's overall utility for alternative primary care facilities [23]. One group of studies assumed optimal-choice, i.e., each client will visit the facility that is optimal with respect to her preferences. Many authors, e.g., Berman [26], Verter and Lapierre [21], Wang et al. [31], and Berman et al. [27], simply assumed that clients patronize the closest facility, whereas Parker and Srinivasan assumed that clients choose the facility with the maximum utility [23]. A second group of studies considered probabilistic-choice, i.e., each client's facility choice is based on a probability distribution, which is generated from the attractiveness and proximity of each facility. Marianov et al. proposed a facility location problem with congestion using a probabilistic-choice model to represent client allocation behavior [32]. Our approach is more practical as we assume that each individual would patronize the healthcare facility that has the minimal expected gross time (which comprises the travel time plus the expected time clients spend at the facility) rather than the closest facility [21]. Consequently, we also assume that the expected number of participants decreases with an increase in the expected gross time, rather than the distance to be traveled. Despite the commonality of these modeling constructs, there are significant differences between our work and those that have been done in the past and are summarized as follows: first of all, Parker and Srinivasan [23] aimed at maximizing total clients' utilities (benefits), whereas our objective is to minimize establishment and staffing costs and the average gross time. Second, they represented a client's overall utility via a weighted linear function of several facility attributes (including, the type of facility, travel time, waiting time, the time to get an appointment, etc.), whereas we only considered the expected total time. Finally, Parker and Srinivasan presented a model for waiting time as a linear function of the number of clients, whereas we used the steady-state expression for an M/M/1 queue in representing the total time spent at a facility as a function of the arrival and service rates. It is perhaps due to these differences that our solution methodology and results are quite different from that of Parker and Srinivasan. More specifically, they do not observe the determination of equilibrium facility–client allocation sets (in which none of the clients is willing to change the facility they patronize according to the allocation). Zhang et al. provided a methodology for designing a network of preventive healthcare facilities so as to maximize participation. The number of facilities to be established and the location of each facility were the main determinants of the configuration of a healthcare facility network. They used the total (travel, waiting and service) time required for receiving the preventive service as a proxy for accessibility of a healthcare facility. Four heuristics were compared in terms of accuracy and computational requirements [22]. In another paper, Zhang et al. presented a methodology for designing a network of preventive healthcare facilities to improve its accessibility to potential clients, and thus maximize participation in preventive healthcare programs [33]. The problem was formulated as a mathematical program with equilibrium constraints. A Tabu search procedure was developed to solve the upper level problem. Gu et al. presented a new methodology for solving the PHCFL problem. They defined a new accessibility measurement that combines the two-step floating catchment area method, distance factor, and the Huff-based competitive model in order to capture the characteristics of preventive health care services. They formulated the PHCFL problem based on the new accessibility measurement as a bi-objective model based on efficiency and coverage and solved it using the Interchange algorithm [5]. Zhang et al. studied the impact of client choice behavior on the configuration of a preventive care facility network and the resulting level of participation. They presented two alternative models, named the "probabilistic-choice model" and the "optimal-choice model". Both models were formulated as a mixed-integer program. They proposed a probabilistic search algorithm and a genetic algorithm to solve the problems [33]. Afshari and Peng provided an overview of methods and challenges for decision-making of the healthcare facility location to ensure an optimal solution. They suggested answers for defined questions. Challenges were discussed in detail for modeling and applications of healthcare facility location problems. Their paper can be used as a methodological guide for the location or relocation of healthcare facilities [4]. Vidyarthi and Kuzgunkaya studied the impact of system optimal choice on the design of the preventive healthcare facility network under congestion. They presented a model that simultaneously determines the location and the size of the facilities as well as the allocation of clients to these facilities to minimize the weighted sum of the total travel time and the congestion associated with waiting and service delay at the facilities. The problem was set up as a network of spatially distributed M/G/1 queues and formulated as a nonlinear mixed integer program, and a cutting plane algorithm-based exact solution approach was presented [34]. Davari et al. developed a fuzzy bi-objective model with budget constraints of the problem [35]. A modified version of the model was introduced by modeling the attractiveness by means of fuzzy triangular numbers and treating the budget constraint as a soft constraint. Two solution methodologies, namely fuzzy goal programming and fuzzy chance-constrained optimization, were proposed as solutions. Both the original and modified models were solved within the framework of a case study in Istanbul, Turkey. Regarding uncertainty in network design, Shishehbori developed a mixed integer non-linear programming formulation to model the combined facility location network design problem with unreliable facilities. He considered different costs including facility location, link construction/improvement, and transportation costs as well as the maximum allowable failure cost of the system in the mathematical formulation [6]. Rohaninejad et al. addressed a reliable facility location problem considering facility capacity constraints. In reliable facility location problem, some facilities may become unavailable. Since failure of facilities could lead to disruptions in facility location decisions, their approach was an attempt to reduce the impact of such disruptions. A novel mixed-integer nonlinear programming model was presented and two different heuristic procedures were developed [36]. Nasiri et al. declared that location of hubs and allocation of demands to them was of high importance in the network design. The most important objective of these models was to minimize the cost, but the importance of path reliability was also considered. They proposed a P-center hub location model with full interconnection among hubs, while there existed different paths between origins and destinations. They determined a reliable path that had lower cost. They utilized Cuckoo Optimization Algorithm in order to solve the problem for large instances [37]. Hosseini-Motlagh addressed a mixed integer linear programming model for blood supply chain network design with the need for making both strategic and tactical decisions throughout multiple planning periods. They used robust programming approach to deal with inherent randomness in parameters. They employed two criteria: the mean and standard deviation of constraint violations under a number of random realizations to assess the performance of both the proposed robust and deterministic models [38]. Taylor et al. presented the continued development of a Safety Risk Assessment (SRA) toolkit to be used proactively during the design of healthcare facility projects [39]. Following content development, the tool was tested at three project sites and through hypothetical scenarios in an interactive testing process engaging expert panels. The testing revealed tactical considerations (content clarity, redundancy, etc.) and strategic aspects (themes related to use) for finalizing the tool. Davari et al. addressed the problem of designing a preventive health care network considering impatient clients and budget constraints [1]. The objective of their model was to maximize the accessibility of services to people. They formulated the problem as a mixed-integer programming problem with budget constraints and congestion considerations. An efficient variable neighborhood search procedure was proposed and computational experiments were performed on a large set of instances. Krohn et al. considered clients' utility function to include variables denoting the waiting time for an appointment and the quality of care. Both variables were defined as a function of a facility's utilization that yields a mixed integer non-linear model formulation. They assumed that the waiting time for an appointment could be considered categorical. The minimum quantity requirement was considered as a categorical variable. They illustrated that the problem can be solved optimally within an acceptable time by applying GAMS /CPLEX to our instances based on both artificial data as well as in the context of a case study based on empirical data [40]. Javanmardi et al. developed a service network design model of preventive healthcare facilities with the principal objective of maximizing participation in the offered services. They considered utility constraints and incorporated demand elasticity of customers due to travel distance and congestion delays, optimal number, locations and capacities of facilities as well as customer assignment to facilities were determined. They solved linearized model by developing an exact algorithm [3]. Ahmadi-Javid et al. declared that the lack of a comprehensive review in the last decade was a serious shortcoming in the literature of healthcare facility location (HCF). They presented a framework to classify different types of non-emergency and emergency HCFs in terms of location management, and reviewed the literature based on the framework. They classified researches on HCF considering criteria such as uncertainty, multi-period setting, particular input/setting, objective function, decision variable, constraint, basic discrete location problem, mathematical modeling approach, solution method, and case study inclusion [41]. Verjan et al. worked on Home Health-Care (HHC), a slowly evolving concept in recent decades. Their idea was to reduce pressure on inpatient hospital beds by providing care to patients at their homes. HHC centers could undertake more complex care such as end-of-life care, chemotherapy, and rehabilitation. They accomplished two main objectives: (i) design a home health-care network by locating HHC centers across a territory, taking into account medical demand and costs of resources and facilities; (ii) optimally manage the activities of HHC centers by deciding on the outsourcing of critical processes for patient care [42].

3. Problem Definition

In this paper, a novel bi-objective preventive healthcare facility network design model within M/M/1 queuing framework is developed. Let $G=(N,L)$ be a network with a set of nodes N ($|N|=n$) and a set of links L. The nodes represent the neighborhoods of a city or some population zones, and the links are the main transportation arteries. The fraction of clients residing at node i is denoted by h_i, $i\in N$. We assume that the number of clients who requires medical service over the entire network is Poisson distributed with a rate of λ per unit of time, and thus $i\in N$ from each node i at a rate λh_i. We assume that there is a finite set of potential locations ($X\in N$) in G for the facilities. We also assume that a single service team in facility located at point j can provide an average of μ_j services per unit of time, $j\in X$; see [43]. We further assume that the service time is exponentially distributed. Therefore, each facility is an M/M/1 queue. We denote the number of technology levels for each potential location of the facility by $k\in M$ ($|M|=m$). This way, parameter μ_j changes to μ_{jk} for each potential location and each level of technology. For the ease of exposition, we also assume that $\mu_{jk}=\mu_k$, $j\in X$, although this assumption can easily be relaxed within the context of our model. Since our paper focuses on a non-appointment system, we can assume the arrival times to the facility to be randomly distributed and the problem as an M/M/1 queue. In a pure appointment-based system for preventive healthcare, the appointments are very often determined for the convenience of both the service provider and the client. Consequently, the arrival times to the facility can be viewed as deterministic, and hence require the use of D/G/1 (or D/M/1) queues. In this approach, however, there is no closed-form expression for the expected waiting time for such queues. It is to be mentioned that a hybrid system which allows both appointment and non-appointment schemes is a very complicated queuing system and beyond the scope of this study. We denote by $\bar{T}_{ij}$ the average total time that individuals from node i spend in order to receive service at facility located at point $j\in X$.

$$\bar{T}_{ij} = t_{ij} + \sum_{k=1}^{m} \bar{W}_{jk} z_{jk} \tag{1}$$

Average total time $\bar{T}_{ij}$ comprises two components: (i) The travel time from node i to facility located at point j through the shortest path denoted by t_{ij}; (ii) The average time clients spend at the facility with the special level of technology including waiting and receiving service, denoted by $\bar{W}_{jk}$. In our model, the assumptions are as follows:

- Number of clients requiring medical service over the entire network is Poisson distributed with a rate of λ per unit of time, and thus from each node i at a rate λh_i, $i\in N$
- There is a finite set of potential locations ($X\in N$) in G for the facilities
- There is a single service team in facility located at point j that can provide an average of μ_j services per unit of time, $j\in X$.
- The service time is exponentially randomly distributed. Therefore, each facility is an M/M/1 queue.
- The number of technology levels for each potential location of facility is denoted by $k\in M$ ($|M|=m$).
- Service rate μ_{jk} is both for each potential location and each level of technology. For the ease of exposition, we also assume that $\mu_{jk}=\mu_k$, $j\in X$, although this assumption can easily be relaxed within the context of our model.

- All individuals from the same node request service from the same facility
- In the long run, the clients will gather sufficient information about the total time required to obtain preventive healthcare services at the facilities in their vicinity, although each client may visit these facilities, infrequently.

The fraction of clients from node i who request service from facility j, denoted by a_{ij}, is a decreasing function of the expected travel time.

$$a_{ij} = \begin{cases} A_{ij} - \gamma(t_{ij}) & if\ t_{ij} < \frac{A_{ij}}{\gamma} \\ 0 & otherwise \end{cases} \quad i \in N, j \in X, k \in M. \tag{2}$$

where A_{ij} is the fraction of clients from node i who would visit facility located at point j when $\overline{T}_{ij} = 0$, i.e., the intercept of the demand decay function, and γ is the slope of the demand decay function. In addition, λj is the rate of clients requesting service from node j. Then,

$$\lambda_j = \lambda \sum_{i=1}^{n} h_i a_{ij} \quad , j \in X. \tag{3}$$

Since the system is an M/M/1 queue,

$$\overline{W}_{jk} = \frac{1}{\mu_k - \lambda_j} \quad j \in X \ , \ k \in M. \tag{4}$$

The objectives of our problem are to find the optimal set of locations $j \in X$ so as to minimize establishment and staffing costs and the average total time. To formulate the problem as a mathematical program, our model includes three decision variables, defined as follows:

$$x_{ij} = \begin{cases} 1 & \text{if clients from node i require service from facility located at point j,} \\ 0 & \text{otherwise.} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if facility is located at node j,} \\ 0 & \text{otherwise.} \end{cases}$$

$$z_{jk} = \begin{cases} 1 & \text{if facility located at point } j \text{, use the level of technology } k, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, the proposed mathematical model is presented as follows:

$$\min \sum_{j \in X} H_j y_j + \sum_{j \in X} \sum_{k=1}^{m} \frac{C_k + C'_k}{\mu_k} z_{jk} \lambda_j \tag{5}$$

$$\min \sum_{i=1}^{n} \sum_{j \in X} \overline{T}_{ij} x_{ij} = \sum_{i=1}^{n} \sum_{j \in X} (t_{ij} + \sum_{k=1}^{m} \overline{W}_{jk} z_{jk}) x_{ij} \tag{6}$$

S.t.

$$\sum_{j \in X} x_{ij} = 1 \qquad i \in N \tag{7}$$

$$x_{ij} \le y_j \qquad i \in N \qquad j \in X \tag{8}$$

$$\sum_{j \in X} H_j y_j \le R_{\max} \tag{9}$$

$$\sum_{k=1}^{m} z_{jk} = y_j \qquad j \in \mathrm{X} \tag{10}$$

$$x_{ij}(t_{ij} + \sum_{k=1}^{m} \overline{W}_{jk} z_{jk}) \le t_{ij} + \sum_{k=1}^{m} \overline{W}_{jk} z_{jk} + M(1 - y_j) \qquad i \in N \quad j \in \mathrm{X} \quad k = 1,2,\ldots,m \tag{11}$$

$$\lambda \sum_{i=1}^{n} h_i a_{ij} x_{ij} \le \sum_{k=1}^{m} \mu_k z_{jk} \qquad j \in \mathrm{X} \tag{12}$$

$$a_{ij} x_{ij} \ge 0 \qquad i \in N \qquad j \in X \tag{13}$$

$$x_{ij}, y_j, z_{jk} = 0,1 \qquad i \in N \quad j \in \mathrm{X} \quad k = 1,2,\ldots,m \tag{14}$$

Where,

$$\overline{W}_{jk} = \frac{1}{\mu_k - \lambda \sum_{i=1}^{n} h_i a_{ij} x_{ij}} \quad j \in X \quad k = 1, 2, \ldots, m$$

$$a_{ij} = A_{ij} - \gamma t_{ij} \quad i \in N \quad j \in \mathrm{X}$$

Constraints (5) and (6) are respectively objective functions wherein the first one defines to minimize establishment and staffing costs, and second one is used to minimize the average total time. Equation (7) ensures that each node is serviced by one facility. Constraint (8) guarantees that clients can require service only from open facilities. Constraint (9) specifies the maximum amount of establishment costs defined by R_{max}. Equation (10) defines that only one technology can be used in each facility. Constraint (11), where M represents a big number, stipulates that clients choose the facility that has minimum expected total time. Constraint (12) indicates the stability of the queue and constraint (13) forbids negative a_{ij}. Constraint (14) indicates binary nature of the decision variables.

As mentioned earlier, since the model is a constrained nonlinear mixed integer programming type, and that exact methods are inefficient to solve it, in the next section, three meta-heuristic algorithms are proposed to find near–optimal solutions.

4. Solving Methodology

Since the proposed mathematical model of the problem at hand is a constrained non-linear integer programming (NLIP) type that is NP-hard, an exact solution is hard (if not impossible) to obtain [44]. In this section, three multi-objective evolutionary algorithms (MOEAs), called NSGA-II, NRGA, and MOSA, successfully applied to complex problems by researchers, are developed to solve the proposed mode.

4.1- Fundamental concept of multi-objective algorithms

Multi-objective evolutionary algorithms (MOEA) are new kinds of the meta-heuristic algorithms that can be defined as the process of simultaneously optimizing two or more conflicting objectives. In multi-objective optimization problems, a vector of decision variables optimizes a vector of objective functions. In other words, from the presence of a number of objectives, a set of optimal solution, namely Pareto optimal solution, is obtained rather than a single optimal solution. Vilfredo Pareto proposed the concept of Pareto optimal solution in 1986.

In the proposed model, there are three conflicting objectives, and in order to generate the Pareto optimal solution, the following must be performed:

$$\min \ \vec{f}(\vec{x}) = \left\{ f_1(\vec{x}), \ \ldots, f_k(\vec{x}) \right\} \tag{15}$$

$$g_i(\vec{x}) \le 0 \ , i = 1, \ 2, \ \ldots, M \tag{16}$$

$$h_j(\vec{x}) = 0 \ , j = 1, \ 2, \ \ldots, P \tag{17}$$

$$\vec{x} \in S$$

where $\vec{x} = \left(x_1, x_2, \ldots, x_n\right)^T$ are decision variables, and $f_i = R^n \to R \ \ (i = 1, \ 2, \ \ldots, \ k)$ are objective functions. Furthermore, $g_i \ , h_j : R^n \to R \ \ (i = 1, \ 2, \ \ldots, \ m; \ j = 1, \ 2, \ \ldots, \ p)$ are inequality and quality constraints, respectively. Now, solution x_1 dominates solution x_2 if:

1) $f_i(\overrightarrow{x_1}) \le f_i(\overrightarrow{x_2}), \ \ \forall i = 1, \ 2, \ \ldots, \ k$
2) $\exists i \in \{1, \ 2, \ \ldots, \ k\} : f_i(\overrightarrow{x_1}) < f_i(\overrightarrow{x_2})$

Under these circumstances, a set of solutions that cannot dominate each other is called Pareto solutions set or Pareto front. Therefore, the objective is to obtain Pareto optimal front. To obtain Pareto optimal front, two main characteristics should be achieved: (I) good convergence of the Pareto front; (II) good diversity within the solutions of the Pareto front.

4.2- The NSGA-II

Srinivas and Deb introduced Non-dominated Sorting Genetic Algorithm (NSGA) [45]. NSGA uses Goldberg’s domination criterion to assign ranks to the solutions. Fitness sharing also is utilized in NSGA to control diversity of solutions in the search space. Performance of NSGA highly depends on the parameters of the fitness sharing and other parameters used in the structure of the algorithm. Hence, Deb et al. proposed an extended version of NSGA, namely NSGA-II [46-47].

Ranking of the solutions is performed by means of Goldberg’s domination criterion. However, as stated in the original reference [46], the complexity order of NSGA-II is reduced by a factor of N according to its predecessor (NSGA), where N indicates the number of solutions (population size). NSGA-II uses a fast non-dominated sorting, with the complexity of O (MN2), to assign the ranks of individuals in the population with size N, where the multi-objective optimization problem has M objective functions. Diversity of solutions is controlled by Crowding Distance in NSGA-II. Crowding distance is defined for solutions of the same rank. The shorter the crowding distance, the more crowded the area, where the solution is in, and vice versa. Hence, when selecting between two solutions with the same rank, the one with the higher crowding distance is preferable. The evaluation process in NSGA-II is shown in Fig. 2.

Mechanism of evolution in NSGA-II is illustrated in Fig. 3, where P_t indicates the main population at iteration t. A mating pool is created and filled, applying the binary tournament selection rule to the main population. In binary tournament selection rule, first, two solutions are selected randomly from the population. The one with lower rank is selected if the solutions are from different ranks. If the solutions have identical rank, the one with higher crowding distance is selected. Crossover and mutation operations are applied to the solutions in mating pool to create a new population Q_t. Main population and new population are merged to create a larger population, R_t. Fast Non-dominated sorting is performed and the solutions in R_t are sorted in several fronts. To create the main population of next iteration P_{t+1}, with the same size as P_t, it is necessary to perform a selection operation. To do so, the fronts are added to P_{t+1}, in increasing order of ranks, until the capacity of P_{t+1} is not exceeded. If without a front, P_{t+1} has fewer elements than P_t, and together with a front it has more elements, the front must be selected partially. To perform a partial selection, the elements of the front are sorted in a decreasing order of crowding distances, and the elements of next iteration are selected from top of the front.

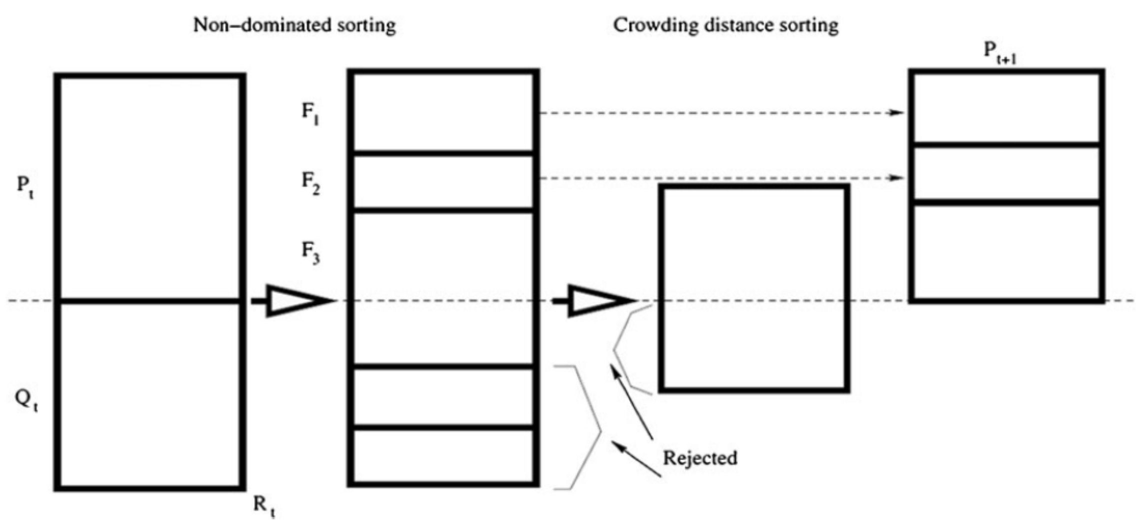

Fig. 2. Evaluation process in NSGA-II

4.2.1- Coding and decoding process

In order to increase the feasibility of the chromosomes in satisfying more constraints, a new type of chromosome is proposed in this research for coding the solution. The coding process takes place in two steps of encoding and decoding, details of which are described in the following two subsections.

The numbers of required facilities associated with the allocation of the customers to the facilities together with the levels of technology associated with the allocation of technology to facilities are decision variables that must be considered in chromosomes. The following four steps describe pertinent portions of a chromosome by which some constraints are satisfied based on the values of these decision variables.

(I) The customer nodes are coded in the first part of the chromosome using a 1×N vector. Each element of this vector contains a random number between zero and one.

(II) The facility nodes are coded in the second part of the chromosome using a 1×M vector. Similarly, each element of this vector includes a random number between zero and one.

(III) The third part of the chromosome consists of a random number (*v*) which is between one and the maximum number of potential facilities that can be on-duty (M).

(IV) The levels of technology nodes are coded in the fourth part of the chromosome using a 1×*v* vector. In a similar manner, each element of this vector includes a random number between one and the maximum number of levels of technology (K).

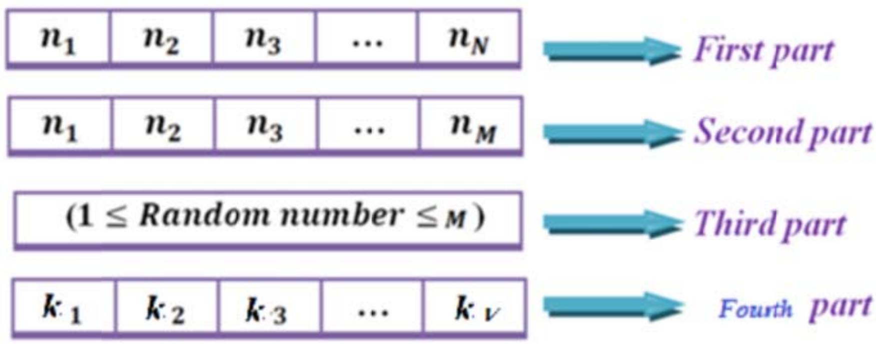

Fig. 3. Solution representation

The decoding process that comes after chromosome representation is one of the most important steps in meta-heuristic algorithm. The parentheses containing numbers in different boxes of this flowchart correspond to step numbers.

In order to better illustrate the coding process consisting of encoding and decoding schemes, consider a numerical example in which N = 5, M = 5, and K = 3. Then, apply the following steps (1)-(8) to both encode and decode the chromosomes in the following manner:

(1) Regarding the third part of a chromosome, a random number (v) between one and five (the maximum number of potential facilities: M), e.g., v = 3, is generated.

(2) Based on the second part of a chromosome, vector m with five genes is generated in which each gene contains a number between zero and one. Fig. 4 shows m vector.

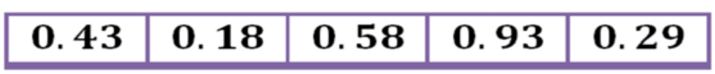

Fig. 4. Generated vector *m*

(3) Sort the genes of vector *m* in an ascending order while reserving the positions as shown in Fig. 5.

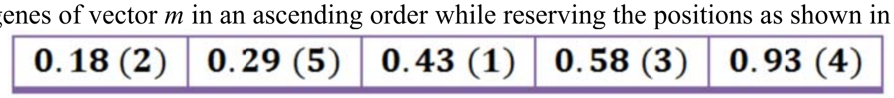

Fig. 5. Sorted vector *m*

(4) As illustrated in Fig. 6, the first three (*v*) genes of the sorted vector are chosen to be open facilities.

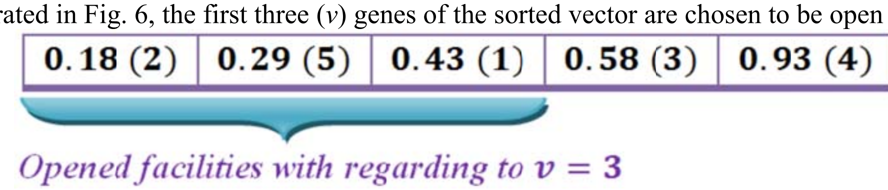

Fig. 6. Opened facilities

(5) This step reports the facilities that are selected for the assignment process. The position number of the potential facilities before the sorting process represents the selected facilities. Fig. 7 illustrates this step.

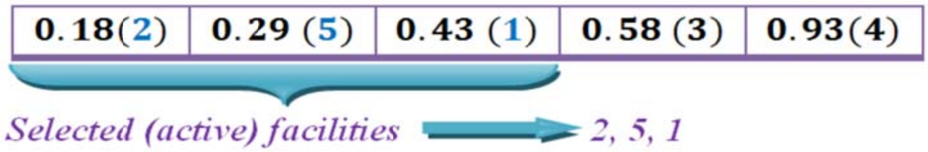

Fig. 7. Solution representation

(6) According to the first part of a chromosome, vector *n* with five genes is generated in which each gene contains a number between zero and one (Fig. 8).

0.65	0.32	0.19	0.09	0.73

Fig. 8. Number of selected facilities

(7) Calculate for *H* vector using $H_i = \left(\lfloor v \times n_i \rfloor + 1\right)$, as shown in Fig. 9.

2	1	1	1	3

Fig. 9. Generated vector *n*

(8) In the final step, based on the fourth part of a chromosome, vector *k* with three (the maximum number of the selected facilities: *v*) genes is generated in which each gene contains number between one and three (the maximum number of levels of technology: K). Fig. 10 shows generated vector *k*.

2	1	3

Fig. 10. Vector *K*

Each cell of the fourth vector will be zero if the corresponding cell of the second vector is zero. This implies that if a facility node is not selected in the second vector, no server can be on-duty in the fourth vector.

The objective-function evaluation step is performed after the decoding process. However, since some constraints are likely to be violated, they are penalized using the method given in [49]. In other words, infeasible solutions are fined using Eq. (18)

$$P(x) = M \times Max\left\{\left(\frac{g(x)}{b} - 1\right), 0\right\} \tag{18}$$

where *M*, g(x), P(x), and f(x) represent a large number, the constraint under consideration, the penalty value, and the objective function value of chromosome x, respectively. In this equation, which is designed for a constraint like g(x) ≤ b, more violations receive bigger penalties. Moreover, penalty values are considered for all of the three objective functions through an additive function given in Eq. (19).

$$F(x) = \begin{cases} f(x) & ; x \in \textit{feasible Region} \\ f(x) + P(x) & ; x \notin \textit{feasible Region} \end{cases} \tag{19}$$

4.2.2- Selection method and elitism

Two operators, called fast non-dominated sorting (FNDS) and crowding distance (CD), are implemented for ranking the population. FNDS assigns the ranks to individuals of the population according to concept of domination wherein a lower value shows a better rank. CD is then calculated for the solutions with the same rank and estimates density of solutions which are laid surrounding a particular solution in the population. In fact, CD is used for controlling the diversity of the solutions and a higher value shows a better solution that is laid in a less crowded area. Details of the calculations of these operators are shown in [46, 48]. Next, binary tournament selection method is used to create the mating pool by randomly selecting two solutions from the population first. When the solutions are from different fronts, the one with lower rank is chosen; otherwise, the one having a higher CD would be selected.

The Elitism process of the NSGAII can also be seen in the above Fig. 2, wherein R_t denotes the main population within the t_{th} iteration. After creating a mating pool by the means of the tournament selection, crossover and mutation operators are applied to the solutions of the mating pool to create a new population Q_t. Then, the main and new population are combined together to create Rt. Subsequently, FNDS is executed and the solutions of R_t sorted in several fronts. Selection operator is again used for creating the main population of the next iteration P_{t+1}, with the same size as Rt. While the capacity of P_{t+1} is not exceeded, the fronts are added to P_{t+1}, in the increasing order of front ranks. In situation without a front, P_{t+1} has fewer members than R_t; in a situation with a front, it has more members, and so the solutions must be selected partially. For partial selection, the members of the front are sorted in the decreasing order of crowding distances, and elements of next iteration are selected from top of the front.

4.2.3- The crossover operator

The following steps demonstrate the crossover operation of this research:

(I) At least, one of the three parts of a chromosome is considered.

(II) Regarding the crossover probability (*Pc*), a number of chromosomes are randomly selected to generate offspring. The number of chromosomes for carrying out the crossover operator is obtained by $Pc \times nPop$.

(III) A continuous crossover operator is implemented in which a random vector (θ) is first generated and the offspring is generated using Eq. (20) [50].

$$\begin{aligned} offspring_1 &= \beta \times parent1 + (1-\beta) \times parent2 \\ offspring_2 &= \beta \times parent2 + (1-\beta) \times parent1 \end{aligned} \tag{20}$$

To illustrate this operation, suppose that a random vector θ with a dimension equal to the size of the selected part (say the second part) of the chosen chromosome (parent) is generated. Then, offspring is obtained using Eq. (20) as shown in Fig. 11. We note that in order to avoid generating infeasible offspring, each exchange is examined to assure feasibility of the generated offspring.

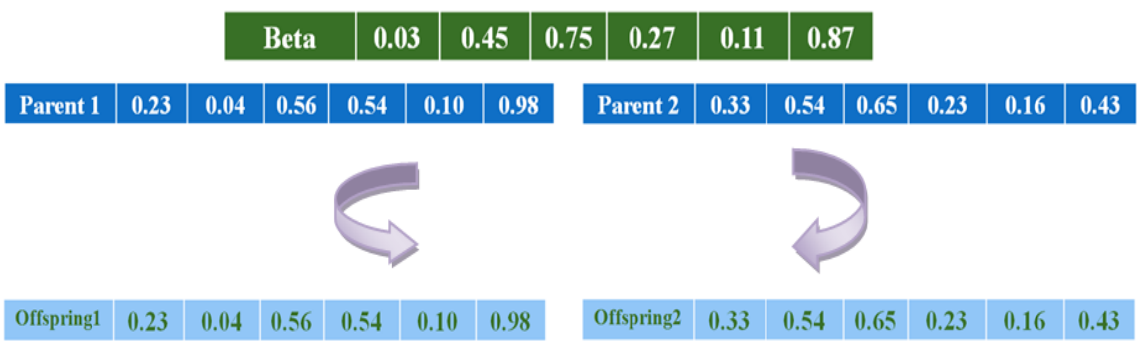

Fig. 11. An example of the crossover operator

4.2.4- The mutation operator

Mutation operator alters a certain percentage of the bits in the list of chromosomes and keeps algorithm from converging too fast before sampling the entire cost surface [48]. The solution spaces that are not discovered by the crossover operator are found using the mutation operator. The steps involved in the mutation operation of this research at each iteration are as follows:

(I) At least, one of the three parts of a chromosome is considered.

(II) With the mutation probability (*Pm)*, a number of chromosomes are randomly selected to generate offspring. This number is obtained by $Pm \times nPop$.

(III) The swap mutation is considered for mutation implementation [50]. In the swap mutation, two positions are randomly selected to swap with each other. Fig. 12 illustrates this operation.

Once again, to avoid infeasible offspring, each exchange is examined upon constraints to assure generating feasible offspring.

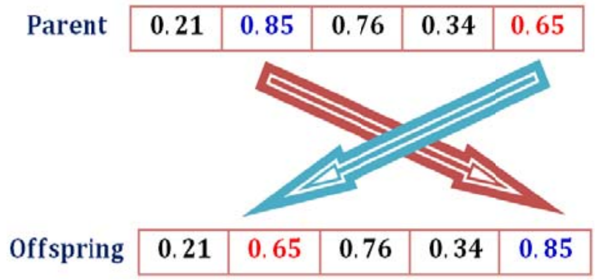

Fig. 12. An example of the mutation operator

4.3- The NRGA

Another popular MOEA, which is utilized in this paper, is NRGA [51]. The main difference between NSGAII and NRGA is their selection strategy. Instead of binary tournament selection, NRGA utilizes roulette wheel selection. In NRGA, after calculating FNDSs and CDs of all individuals, two tires of rank-based roulette wheel selection are used. One of these tires is used for selecting front (based on ranks) and the other one is used for selecting solutions from the front (based on CDs). This way, a front with higher rank has the higher probability to be selected. In addition, in selected fronts, solutions with higher CD are more likely to be chosen. Therefore, the only change for NRGA is in the selection process in which the roulette wheel selection is used instead of binary tournament selection.

4.4- The MOSA

Simulated annealing (SA) was first introduced by Kirkpatrick et al. to obtain near-optimum solutions of optimization models that are hard to solve using conventional procedures [52]. Since then, several authors have employed SA in various optimization problems. SA is a general random search algorithm based on stochastic mechanism of physical annealing process in metallurgy. Generally, the objective value of a solution is equivalent to the internal energy state. The steps involved in the developed SA of this research are explained in the following subsections.

4.4.1- Initialization

In this step, the input parameters of SA are initialized. The parameters are: (1) The initial temperature T_0 which is the temperature at the beginning of each iteration, (2) The final temperature T_h which is the temperature at the end of each iteration, (3) The number of iteration in each temperature *nIt*, and (4) The temperature reduction rate β. Temperature at iteration h, T_h, is obtained using Eq. (21) [52].

$$T_h = \beta \times T_{h-1}; h > 2 \text{ }, 0 < \beta < 1 \tag{21}$$

4.4.2- The coding process

As mentioned in "The developed coding process", to enhance feasibility of solutions and satisfy more constraints, a new type of coding process that includes encoding and decoding schemes is proposed. These schemes for SA are similar to the ones described for NSGA-II.

4.4.3- Main loop of the SA

SA starts with a high temperature and randomly chooses initial solution ω_0. The initial value of T_0 acts as a controller parameter of the temperature, and a new solution ω_n within the neighborhood of the current solution ω is calculated at each iteration. If the value fitness function, $f(\omega_n)$, is less than the previous value $f(\omega)$, the new solution is accepted. Otherwise, in order to escape from the local optimal solution, the new solution is accepted with a probability (*Probabilty*$_{SA}$) derived from Eq. (22) [52].

$$\text{Probability}_{SA} = e^{-\frac{\Delta}{T}} \quad ; \quad \Delta = \frac{f(\omega_n) - f(\omega)}{f(\omega_n)} \times 100 \tag{22}$$

This process is repeated until the desired state of the algorithm is reached.

4.4.4- Neighborhood representation

To represent the neighborhood structure, the proposed mutation operator of NSGA-II, described in "The mutation operator", is utilized to avoid fast convergence of SA.

4.4.5- Multi-objective operators of the MOSA

While the objective function value is used to rank the solutions in a single-objective algorithm, the domination concept is utilized for ranking in Pareto-based multi-objective algorithms. In the NSGA-II algorithm [46, 48], the fast non-dominated sorting (FNDS) operator was employed for inserting the dominance concept by searching the first goal called convergence. Smaller values of FNDS indicate better ranks. To search for the second goal named diversity, another operator named crowding distance (CD) was considered in NSGA-II to estimate the density of similar rank solutions laid surrounding a particular solution. Bigger values of CD show better solutions lying in a less crowded area. Next, a binary tournament selection is performed according to the above two operators, in which if solutions are from different ranks, the one with smaller rank is selected, otherwise, the one having a higher value of CD is used. Fig. 13 illustrates the pseudo code of MOSA algorithm based on the basic operators of a SA algorithm and the described multi-objective operators.

```
Parameter setting: popsize, nMove, num.it,  frontmax
Initialization: Generate initial solutions
Evaluation: Evaluate initial solutions

Perform non-dominate sorting and calculate ranks
Calculate crowding distance (CD)
Sort population according to ranks and CDs

Pt=population
For  it=1: num.it

   for i=1:popsize
      for j=1:nMove
         St(i)= perform  neighborhood structure on the solution i of the population
      end
   end

   Perform non-dominate sorting and calculate ranks (St)
   Calculate crowding distance (CD) (St)
   Sort population according to ranks and CDs (St)

   for i=1:popsize
      if ~Dominates (Pt(i), St (i))
         Qt  (i)= St (i)
     else
        delta=Cost Pt (i)- Cost  St (i)
        p=exp(-delta/T(it))
        if rand<p
           Qt  (i)= St (i)
        end
     end
   end

   Rt = Pt ∪Qt
   Perform non-dominate sorting and calculate ranks (Rt)
   Calculate crowding distance (CD) (Rt)
   Sort population according to ranks and CDs (Rt)

   if  size( Rt )> frontmax
      Pt =Select  frontmax number of the solution
      non-dominate sorting and calculate ranks  ( Pt )
      Calculate crowding distance  ( Pt)
   end
Update T
End
```

Fig. 13. Pseudo code of the MOSA

4.4.6- Stopping criteria

Stopping criteria are a set of conditions such that a good solution is obtained when satisfied. While different criteria are used to stop the algorithms, in this research, algorithms stops when an improvement in the fitness function values for several successive generations is not achieved.

5. Computational Results

This section provides details of the proposed methodology and the performance comparisons of the three meta-heuristic algorithms using a parameter tuning procedure. Initially, some multi-objective performance metrics are introduced.

5.1- Multi-objective metrics

In order to evaluate the performances of the three multi-objective meta-heuristic algorithms, the following five metrics are used:

1. Diversity, to measure the extension of the Pareto front [53].
2. Spacing, to measure the standard deviation of the distances among solutions of the Pareto front [54].
3. Mean ideal distance (MID), to measure the convergence rate of Pareto fronts to a certain point (0,0) [53].

4. Number of found solutions (NOS), to count the number of the Pareto solutions in Pareto optimal front.
5. The CPU time: to measure the amount of time required by the algorithms to reach near-optimum solutions.

5.2- Parameter setting

In the following two subsections, we present parameter settings of both the model and algorithms.

5.2.1- Tuning model parameter

To assess the model, 20 test problems are generated randomly. These problems are categorized based on the number of customers (I), the number of facilities (J), and the maximum number of on-duty servers (K). Each test problem is employed three times and the average solution values are obtained and used for performance evaluations. Tables 1 and 2 contain different values of these parameters and provide some relevant additional information.

Tab. 1. Input of the model

Test Problem Number	I	J	K
1	4	4	2
2	7	7	2
3	9	9	3
4	12	12	3
5	14	14	4
6	17	17	4
7	19	19	4
8	22	22	5
9	25	25	5
10	27	27	6
11	30	30	6
12	32	32	6
13	35	35	7
14	37	37	7
15	39	39	8
16	41	41	8
17	43	43	9
18	45	45	9
19	48	48	10
20	50	50	10

Tab. 2. Input of the parameter

Parameter	Value	Parameter	Value
λ	10	h_i	Uniform(0,1)
μ_i	Uniform(7,10)	H_j	Uniform(200000,300000)
t_{ij}	Uniform(0,1)	C_k	Uniform(10,50)
$C_k^{'}$	Uniform(10,50)	$R_{\max}$	800000
γ	0.55	A_{ij}	0.95

5.2.2- Tuning algorithm parameters

In order to calibrate the parameters of the proposed algorithms, the Taguchi method is utilized in this research. Taguchi method is a fractional factorial experiment (FFE) that is proposed by Taguchi as an efficient alternative to full-factorial experiments [55]. Taguchi method uses orthogonal arrays for setting family of experiences to study a group of decision variables or factors. In this method, factors are categorized into two groups: (1) controllable or signal factors; (2) noise factors. Now, based on the concept of robustness, the method seeks to minimize the effect of noise and determine the optimal level of signal factors. To do so, the signal-to-noise ratio (S/N), which calculates the amount of variation of the response, is implemented. According to the type of the response, the variation calculation in the Taguchi approach is classified into three main groups of (1) smaller-the-better type, (2) nominal-is-best type, and (3) larger-the-better type. Then, the aim of the method is to maximize the S/N ratio. For more details on the Taguchi approach, one can refer to [55-56]. Due to minimization nature of the objective functions of this research, the smaller-the-better type of the response is used. Eq. (23) formulates S/N of this type of response, where Y denotes the response and n shows the number of orthogonal arrays.

$$\left(S/N\right) = -10 \times \log\left(\frac{S(Y^2)}{n}\right) \tag{23}$$

To conduct the Taguchi method more comprehensively, MID metric is considered in this research. As mentioned, in Pareto-based algorithm, one of the main goals is good convergence. Among introduced metrics in Section 5.1, CPU time and MID are the ones that measure the convergence rate of the algorithms.

In order to utilize the Taguchi method, the levels of the factors are first determined in Table 3. As observed, factors are presented in two ways: their actual names along with their coded names. For example, in NSGA-II, A represents nPop. Moreover, three levels are considered for each factor involved in the algorithms. Then, using Minitab Software, the L9 design is used for NSGA-II, NRGA, and MOSA. The orthogonal arrays of these designs along with experimental results are shown presented in Table 4 (for NSGA-II and NRGA) and Table 5 (for MOSA). For each algorithm, the effect plot for S/N ratio is presented in Fig. 14.

The highlighted cells of Table 3 show proper levels of the parameters in all algorithms. For the other algorithms, a similar approach is used, wherein the selected levels of their parameters are also the highlighted ones in Table 3.

Tab. 3. Algorithm parameter ranges along with their level

Solving methodologies	Parameter	Low	Medium	High
$NSGA-II$	P_c	0.6	0.8	0.99
	P_m	0.01	0.2	0.4
	$nIt_{NSGA-II}$	100	300	500
	$nPop_{NSGA-II}$	25	100	200
$NRGA$	P_c	0.6	0.8	0.99
	P_m	0.01	0.2	0.4
	$nIt_{NSGA-II}$	100	300	500
	$nPop_{NSGA-II}$	25	100	200
MOSA	T_f	10	5	1
	T_0	100	200	300
	nIt_{SA}	100	300	500
	β	0.9	0.8	0.7

Tab. 4. Computational results to tune NSGA-II and NRGA

Run order	$NSGA-II$ & $NRGA$ Parameters				MID Measure	
	P_c	P_m	nIt_{GA}	$nPop_{GA}$	NSGA-II	NRGA
1	0.6	0.01	100	25	755971	1963552
2	0.6	0.2	300	100	647170	229678
3	0.6	0.4	500	200	651400	232608
4	0.8	0.01	300	200	282613	556646
5	0.8	0.2	500	25	573698	238247
6	0.8	0.4	100	100	244247	252179
7	0.99	0.01	500	100	211143	740841
8	0.99	0.2	100	200	688902	688349
9	0.99	0.4	300	25	689899	656008

Tab. 5. Computational results to tune MOSA

Run order	$MOSA$ Parameters				MID Measure
	T_f	T_0	nIt_{SA}	β	MOSA
1	10	100	100	0.9	795150
2	10	200	300	0.8	663042
3	10	300	500	0.7	654384
4	5	100	300	0.7	873605
5	5	200	500	0.9	582942
6	5	300	100	0.8	684548
7	1	100	500	0.8	619120

8	1	200	100	0.7	877596
9	1	300	300	0.9	734622

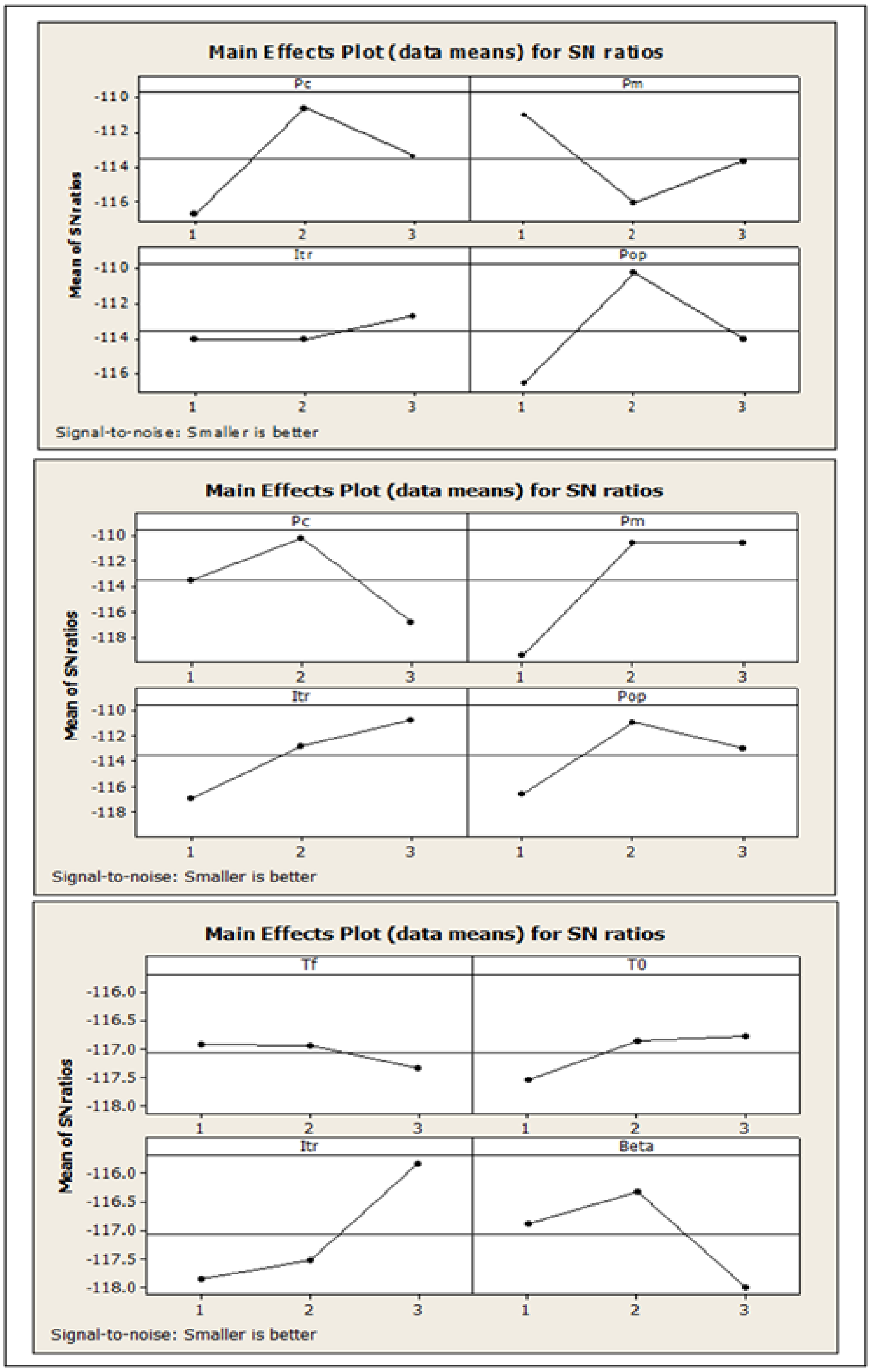

Fig. 14. Taguchi ratios for all proposed algorithms

5.3- Results analysis

In this section, the performances of the proposed tuned multi-objective solving methodologies are evaluated and compared using the multi-objective metrics given in Section 4.1. Tables 6, 7, and 8 contain the computational results of employing the algorithms on the 20 test problems introduced in Section 4.2.1, where "NA" shows that the algorithm cannot find Pareto front in the reported time. For these cases, the metric based on the whole test problems is also plotted in Fig. 15. Moreover, the algorithms are statistically compared based on the properties of their obtained solutions via the analysis of variance method. The P-values of these tests on each metric are summarized in Table 9. To clearly and highlight the results of the tests, for the cases where a significant difference is obtained, Interval plots are shown in Fig. 16.

Tab. 6. Multi-objective metrics obtained for NSGA-II

	Proposed NSGA-II						
	N	*K*	*Diversity*	*NOS*	*Spacing*	*MID*	*CPU Time*
1	4	2	17652	6	7.2175	549616	614
2	7	2	66450	4	0.104	498996	608
3	9	3	442994	9	130604	613094	615
4	12	3	177074	2	0	651352	643
5	14	4	406409	2	0	516649	636
6	17	4	0	1	NaN	289466	653
7	19	4	179018	2	0	623219	629
8	22	5	202702	3	76071	739988	632
9	25	5	161412	3	68171	720135	631
10	27	6	215223	4	16592	761161	611
11	30	6	187613	2	0	689675	647
12	32	6	189548	2	0	745928	657
13	35	7	188086	2	0	737407	640
14	37	7	20038	2	0	242570	628
15	39	8	238085	3	80802	851085	635
16	41	8	224882	3	45518	847165	643
17	43	9	189992	3	76915	854764	644
18	45	9	0	1	NaN	245010	677
19	48	10	219570	3	63897	883127	636
20	50	10	225352	4	12046	830595	632
SUM	-	-	3552100	61	570623.3	12891002	12711

Tab. 7. Multi-objective metrics obtained for NRGA

	Proposed NRGA						
	N	*K*	*Diversity*	*NOS*	*Spacing*	*MID*	*CPU Time*
1	4	2	225456	5	84425	384343	698
2	7	2	0	1	NaN	426051	736
3	9	3	12.9	2	0	488690	719
4	12	3	0	1	NaN	261608	754
5	14	4	175264	3	61479	644425	715
6	17	4	0	1	NaN	235741	766
7	19	4	85795	2	0	652356	729
8	22	5	135464	2	0	611351	737
9	25	5	0	1	NaN	696940	771
10	27	6	303228	4	101673	789287	715
11	30	6	0	1	NaN	245845	761
12	32	6	0	1	NaN	676461	757
13	35	7	166664	2	0	734383	737
14	37	7	0	1	NaN	217202	784
15	39	8	186740	2	0	775834	747
16	41	8	173029	2	0	698443	738
17	43	9	170845	2	0	810799	747
18	45	9	149891	2	0	758514	745
19	48	10	177339	2	0	784194	749
20	50	10	68940	2	0	257215	740
SUM	-	-	2018668	39	247577	11149682	14845

Tab. 8. Multi-objective metrics obtained for MOSA

	Proposed MOSA						
	N	*K*	*Diversity*	*NOS*	*Spacing*	*MID*	*CPU Time*
1	4	2	181279	5	55514	465913	201
2	7	2	455336	7	65886	517690	201
3	9	3	19621	2	0	724960	208
4	12	3	0	1	NaN	1051993	206
5	14	4	0	1	NaN	437662	206
6	17	4	0	1	NaN	908204	215
7	19	4	62318	2	0	280689	206
8	22	5	161940	3	101342	637885	212
9	25	5	79514	2	0	774178	209
10	27	6	55354	2	0	656828	207
11	30	6	161867	4	21648	754252	207
12	32	6	0	1	NaN	4496719	213
13	35	7	81863	2	0	784064	209
14	37	7	0	1	NaN	705376	207
15	39	8	0	1	NaN	1830950	213
16	41	8	0	1	NaN	5987146	213
17	43	9	0	1	NaN	4611612	215
18	45	9	0	1	NaN	2479662	212
19	48	10	0	1	NaN	1332175	210
20	50	10	0	1	NaN	3572368	212
SUM	-	-	1259092	40	244390	33010326	4182

We note that, while in terms of the diversity and NOS metrics, bigger values are desired; for spacing, MID and CPU time, smaller values are better. Therefore, in general, based on the outputs in the last row of Tables 6, 7, and 8, it is clear that NSGA-II shows better performances in terms of diversity and NOS. Meanwhile, for MID metric, NRGA, and for spacing and CPU time, MOSA has better performance. However, when the metrics are statistically compared, Fig. 15 shows that only in terms of spacing, the algorithms have no significant differences. This conclusion is also confirmed at 95% confidence level based on the results given in Table 9. Furthermore, Figs. 15 and 16 support this conclusion as well.

Tab. 9. The P-values of the analysis of variance comparison test

Metric's name	*P-value*	*Test results*
Diversity	0.005	Null hypothesis is rejected
Spacing	0.216	Null hypothesis is not rejected
NOS	0.045	Null hypothesis is rejected
MID	0.001	Null hypothesis is rejected
CPU Time	0.000	Null hypothesis is rejected

In addition, it should be noted that MATLAB Software (Version 7.10.0.499, R2010a) [57] was used to code the proposed meta-heuristic algorithms, and the programs were executed on a 2.53 GHz laptop with 1 GB RAM.

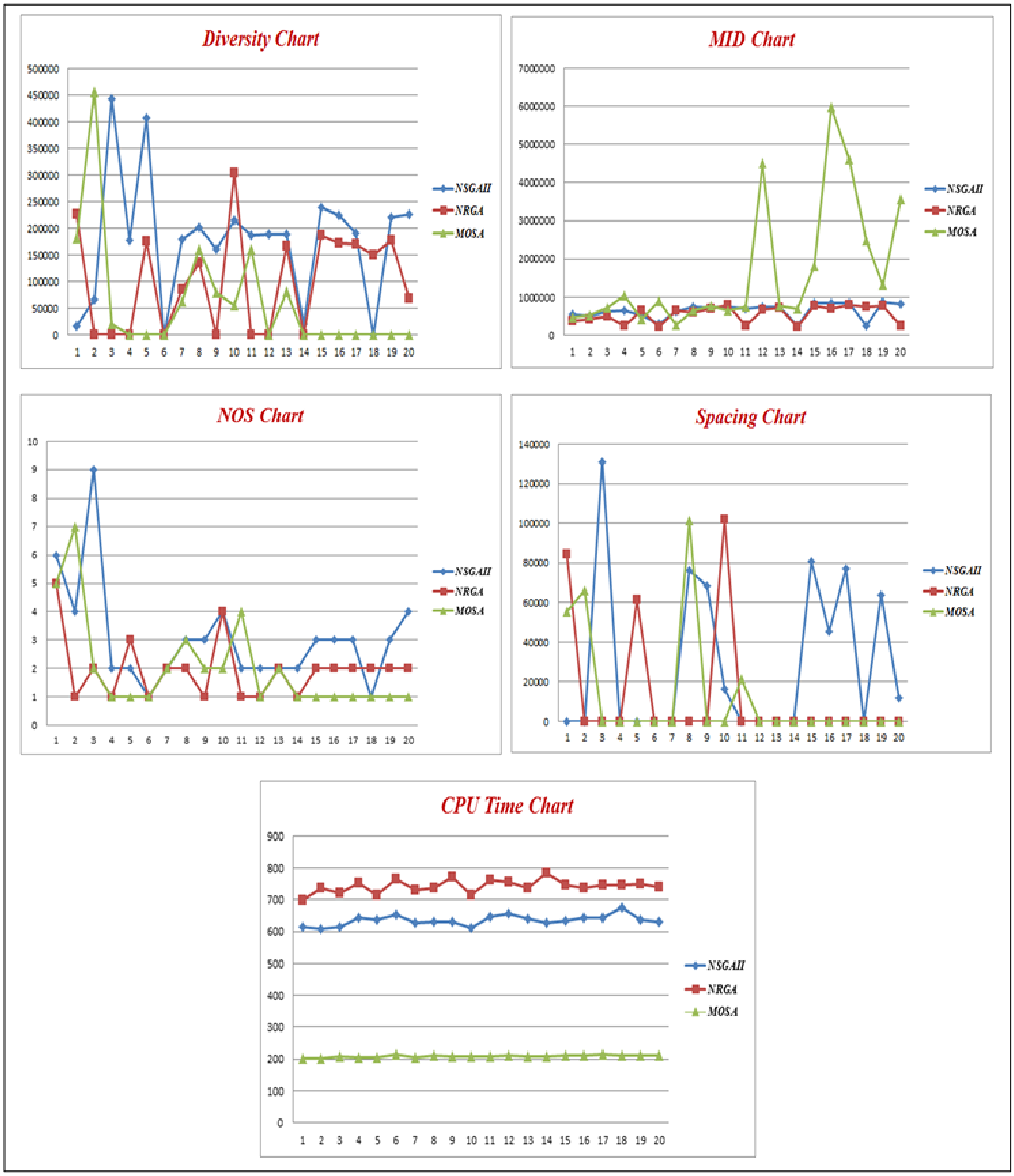

Fig. 15. Detailed comparisons of the algorithms on all test problems

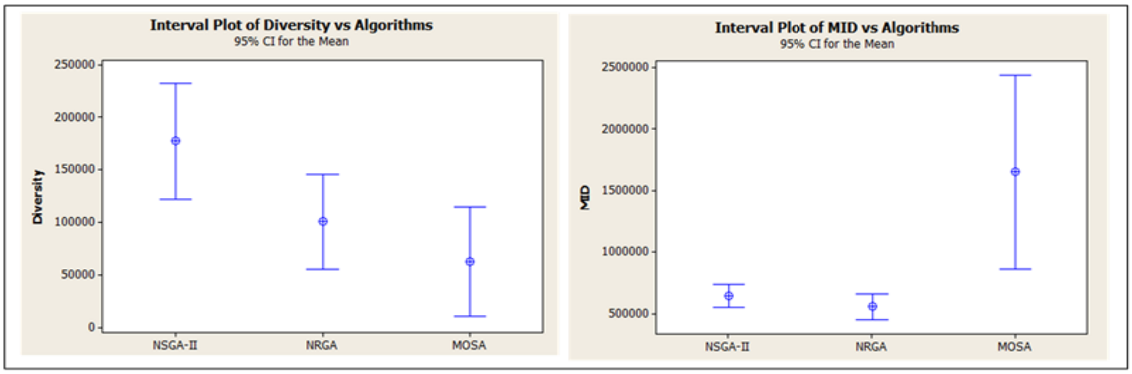

Fig. 16. Box-plot of metrics with significant difference

6. Conclusion and Future Researches

This paper presented a nonlinear programming model for designing optimal preventive healthcare facility networks so as to maximize participation in preventive healthcare system. The proposed model incorporated three important characteristics of the problem, namely elastic demand, congestion, and user-choice environment. Although each facility is an M/M/1 queue and the demand decay function is linear in our model, the proposed methodology can also be used for M/G/1 queues, where waiting time is an increasing function of the arrival rate, and for other monotonically decreasing demand decay functions. Due to highly nonlinear nature of the model, we developed an allocation and four location heuristics. The objective functions minimized travel and waiting time as well as total cost, which included establishment and staffing cost. Subsequently, in view of the fact that PHPs are basically NP-Hard, three parameter tuned Pareto-based multi-objective meta-heuristic algorithms, called NSGA-II, NRGA, and MOSA were proposed to solve the problem. Performance of the proposed algorithms was then statistically compared using 20 randomly generated test problems via five multi-objective metrics. Finally, based on the obtained results, it was shown that:

- Performance of the algorithms is similar based on spacing metric.
- NSGA-II significantly performs better than the NRGA and MOSA in terms of NOS and diversity metrics.
- NRGA performance is much better than the NSGA-II and MOSA in terms of MID metric.
- In terms of CPU time, while NSGA-II and NRGA performances are similar, MOSA outperforms them.

It is clear that by adopting our proposed methodology in the design of a network of preventive healthcare facilities, one can minimize their establishment and staffing costs, as well as the total time required for the service. Considering the fact that such factors are directly related to participation and there are empirical evidences that the convenience of access plays a key role in the participation, we can safely say that its actual implementation could lead to maximum participation. In short, our findings provide managerial insights into maximum preventive healthcare participation achievement strategy by means of total service time reduction, attractive facility proximity and its surroundings on top of cheaper establishment and maintenance cost.

The followings can be considered for some relevant future research:

- Use of other queuing disciplines to model QPHP.
- Adoption of a different QPHP model when customers encounter multi-echelon queuing networks.
- Assignments of numerous service rates to the facilities.
- Employment of a different all-feasible chromosome representation.
- Exploitation of different multi-objective solution methodologies. The demand and service rates can be considered fuzzy inputs to model a $\tilde{M}/\tilde{M}/1$ queuing system.

Acknowledgement

The authors would like to sincerely thank Qazvin Islamic Azad University for their numerous supports in carrying out this study.

References

[1] Davari, S., Kilic, K., Naderi, S., *"A heuristic approach to solve the preventive health care problem with budget and congestion*

constraints," Applied Mathematics and Computation, Vol. 276, 2016, pp. 442-453.

[2] Zhang, Y., Berman, O., Verter, V., *"The impact of client choice on preventive healthcare facility network design,"* OR Spectrum, Vol. 34, Issue 2, 2012, pp. 349-370.

[3] Javanmardi, S., Hosseininasab, H., Mostafaipour, A., *"An Exact Method for Stochastic Maximal Covering Problem of Preventive Healthcare Facilities,"* Journal of Industrial and Systems Engineering, Vol. 10, Issue 1, 2010, pp. 10-23.

[4] Afshari, H., Peng, Q., *"Challenges and Solutions for Location of Healthcare Facilities,"* Industrial Engineering and Management, Vol. 3, Issue 2, 2014, pp. 1-12.

[5] Gu, W., Wang, X., McGregor S.E., *"Optimization of preventive health care facility locations,"* International Journal of Health Geographics, Vol. 9, Issue 17, 2010, pp. 1-16.

[6] Shishehbori, S., *"Study of Facility Location-network Design Problem in Presence of Facility Disruptions: A Case Study,"* International Journal of Engineering, Vol. 28, 2015, pp. 97-108.

[7] Walker, K., *"Current issues in the provision of health care services,"* Journal of Consumer Affairs. Vol. 11, 1977, pp. 52-62.

[8] Health Canada, Mammography, Retrieved June 2014 from: http://www.hc-sc.gc.ca/iyh-vsv/med/mammog_e.html, 2005.

[9] Gornick, M.E., Eggers, P.W., Riley, G.F., *"Associations of race, education, and patterns of preventive service use with stage of cancer at time of diagnosis,"* Health Services Research, Vol. 39, 2004, pp. 1403-1427.

[10] World Health Organization, *"Integrating Prevention into Health Care. Fact.Sheet172,"* Retrieved June 2014 from: http://www.who.int/mediacentre/factsheets/fs172/en/index.html, 2002.

[11] Falkenheimer, S.A., The Adequacy of Preventive Health Care: Does the Health Care Provider Matter? Retrieved June 2014 from: http://www.cbhd.org/resources/healthcare/falkenheimer_2004-09-24.htm, 2004.

[12] Goldsmith, J., *"A radical prescription for hospitals,"* Harvard Business Review, Vol. 67, 1989, pp. 104-111.

[13] Institute of Medicine, Access to Health Care in America, National Academy Press, Washington, DC, 1993.

[14] Zimmerman, S., *"Factors influencing Hispanic participation in prostate cancer screening,"* Oncology Nursing Forum, Vol. 24, 1997, pp. 499-504.

[15] McNoe, B., Richardson, A.K., Elwood, J.M., *"Factors affecting participation in mammography screening,"* New Zealand Medical Journal, Vol. 109, 1996, pp. 359-362.

[16] Facione, N.C., *"Breast cancer screening in relation to access to health services,"* Oncology Nursing Forum, Vol. 26, 1999, pp. 689-696.

[17] US Food and Drug Administration, Mammography quality standards, Center for Devices and Radiological Health Publication, Retrieved June 2014 from: http://www.fda.gov/cdrh/dmqrp.html, 1999.

[18] Daskin, M.S., Dean, L.K., Location of health care facilities. In: Brandeau, M.L., Sainfort, F., Pierskalla, W.P. (Eds.), Operations Research and Health Care: A Handbook of Methods and Applications, Kluwer's International Series, 2004, pp. 43-76.

[19] Berman, O., Krass, D., Facility location problems with stochastic demands and congestion. In: Drezner, Z., Hamacher, H.W. (Eds.), Facility Location: Applications and Theory. Springer, New York, 2002b, pp. 331-373.

[20] Marianov, V., Serra, D., Location problems in the public sector. In: Drezner, Z., Hamacher, H.W. (Eds.), Facility Location: Applications and Theory. Springer, New York, 2002, pp. 119-150.

[21] Verter, V., Lapierre, S.D., *"Location of preventive health care facilities,"* Annals of

Operations Research, Vol. 110, 2002, pp. 123–132.

[22] Zhang, Y., Berman, O., Verter, V., *"Incorporating congestion in preventive healthcare facility network design,"* European Journal of Operational Research, Vol. 198, 2009, pp. 922‑935.

[23] Parker, B.R., Srinivasan, V., *"A consumer preference approach to the planning of rural primary health-care facilities,"* Operations Research, Vol. 24, 1976, pp. 991-1025.

[24] Berman, O., Parkan, C., *"A facility location problem with distance dependent demand,"* Location Science, Vol. 12, 1981, pp. 623-632.

[25] Berman, O., Kaplan, E.H., *"Facility location and capacity planning with delay-dependent demand,"* International Journal of Production Research, Vol. 25, 1987, pp. 1773-1780.

[26] Berman, O., *"The maximizing market size discretionary facility location problem with congestion,"* Socio-Economic Planning Science, Vol. 29, 1995, pp. 39-46.

[27] Berman, O., Drezner, Z., *"Location of congested capacitated facilities with distance sensitive demand,"* IIE Transactions, Vol. 38, 2006, pp. 213-231.

[28] Berman, O., Krass, D., *"The generalized maximal covering location problem,"* Computers and Operations Research, Vol. 29, 2002a, pp. 563-581.

[29] Berman, O., Krass, D., Wang, J., *"Locating service facilities to reduce lost demand,"* IIE Transactions, Vol. 38, 2006, pp. 933–946.

[30] Gerard, K., Shanahan, M., Louviere, J., *"Using stated preference discrete choice modelling to inform healthcare decision-making: A pilot study of breast screening participation,"* Applied Economics, Vol. 35, 2003, pp. 1073-1085.

[31] Wang, Q., Batta, R., Rump, C.M., *"Algorithms for a facility location problem with stochastic customer demand and immobile servers,"* Annals of Operations Research, Vol. 111, 2002, pp. 17–34.

[32] Marianov, V., Rios, M., Icaza, M.J., *"Facility location for market capture when users rank facilities by shorter travel and waiting times,"* European Journal of Operational Research, Vol. 191, 2008, pp. 32-44.

[33] Zhang, Y., Berman, O., Marcotte, P., Verter, V., *"A bilevel model for preventive healthcare facility network design with congestion,"* IIE Transactions, Vol. 42, 2010, pp. 865-880.

[34] Vidyarthi N., Kuzgunkaya, O., *"The impact of directed choice on the design of preventive healthcare facility network under congestion,"* Health Care Management Science, DOI: 10.1007/s10729-014-9274-2, 2014.

[35] Davari, S., Kilic, K., Ertek, G., Fuzzy, *"Bi-objective Preventive Health Care Network Design,"* Health Care Management Science, DOI: 10.1007/s10729-014-9293-z, 2014, pp. 1-15.

[36] Rohaninejad, M., Amiri, A., Bashiri, M., *"Heuristic methods based on MINLP formulation for reliable capacitated facility location problems,"* International Journal of Industrial Engineering & Production Research, Vol. 26, Issue 3, 2015, pp.229-246.

[37] Nasiri, M.M., Shamsi Gamchi, N., Torabi, S.A., *"Cuckoo optimization algorithm for a reliable location-allocation of hubs among the clients,"* International Journal of Industrial Engineering & Production Research, Vol. 27, Issue 4, 2016, pp.309-320.

[38] Hosseini-Motlagh, S.M., Cheraghi, S., Ghatreh Samani, M., *"A robust optimization model for blood supply chain network design,"* International Journal of Industrial Engineering & Production Research, Vol. 27, Issue 4, 2016, pp.425-444.

[39] Taylor E., Quan, X., Joseph, A., *"Testing a tool to support safety in healthcare facility design,"* Procedia Manufacturing, Vol. 3, 2015, pp. 136-143.

[40] Krohn, R., Muller, S., Haase, K., Preventive Health Care Facility Location Planning with

Quality-Conscious Clients, University of Hamburg, 2016, pp. 1-11 (Working Paper).

[41] Ahmadi-Javid, A., Seyedi, P., Syam, S.S., *"A survey of healthcare facility location,"* Computers & Operations Research, Vol. 79, 2017, pp.223-263.

[42] Verjan, C.R., Augusto, V., Xie, X., *"Home health-care design: location and configuration of home health-care centers,"* Operations Research for Health Care, 2017, In Press.

[43] Gunes, E.D., Chick, S.E., Zeynep, A.O., *"Breast cancer screening services: Trade-offs in quality, capacity, outreach, and centralization,"* Health Care Management Science, Vol. 7, 2004, pp. 291-303.

[44] Ehrgott, M., & Gandibleux, X., *"An annotated bibliography of multi-criteria combinatorial optimization,"* OR Spectrum, Vol. 22, 2000, pp. 425-460.

[45] Srinivas, N., Deb, K., *"Multi-objective optimization using non-dominated sorting in genetic algorithms,"* Evolutionary Computation, Vol. 2, Issue 3, 1994, pp. 221–248.

[46] Deb, K., Agrawal, S., Pratap, A., Meyarivan, T., *"A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II,"* In: Proceedings of The Parallel Problem Solving From Nature VI (PPSN-VI) Conference, 2000, pp. 849-858.

[47] Deb., Pratap, A., Agarwal, S., Meyarivan, T., *"A fast and elitist multiobjective genetic algorithm: NSGA-II,"* IEEE Transactions on Evolutionary Computation, Vol. 6, 2002, pp. 182-197.

[48] Deb, K., Multiobjective optimization using evolutionary algorithms. Wiley, Chichester, U.K., 2001.

[49] Yeniay, O., Ankare, B., *"Penalty function methods for constrained optimization with genetic algorithms,"* Mathematical and Computational Application, Vol. 10, 2005, pp. 45-56.

[50] Gross, D., Harris, C. M., Fundamental of queuing theory (3rd ed), New York, NY: Wiley Interscience, 1998.

[51] Al Jadaan, O., Rao, C.R., Rajamani, L., *"Non-dominated ranked genetic algorithm for solving multi-objective optimization problems: NRGA,"* Journal of Theoretical and Applied Information Technology, 2008, pp. 60–67.

[52] Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P., *"Optimization by simulated annealing,"* Science, Vol. 220, 1983, pp. 671-680.

[53] Zitzler, E., Thiele, L., *"Multi-objective optimization using evolutionary algorithms a comparative case study, in: A.E. Eiben, T. Back, M. Schoenauer, H.P. Schwefel (Eds.),"* Fifth International Conference on Parallel Problem Solving from Nature (PPSN-V), Berlin, Germany, 1998. pp. 292-301.

[54] Zitzler, E., Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications. PhD. Thesis, Dissertation ETH No. 13398, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland, 1999.

[55] Peace, G.S., Taguchi Methods, Addison-Wesley Publishing Company, 1993.

[56] Niaki, S.T.A., Ershadi, M.J., *"A parameter-tuned genetic algorithm for statistically constrained economic design of multivariate CUSUM control charts: a Taguchi loss approach,"* International Journal of Systems Science, Vol. 43, 2012, pp. 2275-2287.

[57] MATLAB, Version 7.10.0.499 (2010). The MathWorks, Inc. Protected by U.S. and international patents.

Follow This Article at The Following Site

Roshan K, Seifbarghy M, Pishva D. Multi-objective evolutionary algorithms for a preventive healthcare facility network design. IJIEPR. 2017; 28 (4) :403-427
URL: http://ijiepr.iust.ac.ir/article-1-754-en.html