Journal of Industrial and Systems Engineering
Vol. 13, No. 1, pp. 42-75
Winter (January) 2020

An efficient centralized master echocardiography schedule in a distributed hospital/clinic network

Delaram Chaghazardy[1], Seyed Hessameddin Zegordi[1*], Mohammad Mehdi Sepehri[1], Hassan Aghajani[2]

[1] *Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran*
[2] *Department of Cardiology, Tehran Heart Centre, Tehran University of Medical Sciences, Tehran, Iran*

d.chaghazardy@modares.ac.ir, zegordi@modares.ac.ir, mehdi.sepehri@modares.ac.ir, aghajanih@tums.ac.ir

Abstract

Appointment scheduling systems are applied in a broad variety of healthcare environments to reduce costs and increase quality of services. This study is concerned with the problem of appointment scheduling in a distributed multi-hospital network of echocardiography departments. In this paper, a centralized master schedule is presented to maximize profit margin through maximizing the number of performed echoes and minimizing overtime. Developing such a schedule requires handling shift scheduling and capacity allocation problems simultaneously. Based on real-world settings, a mixed integer linear programming model is proposed for the research problem. Since this model requires a large amount of time and memory to provide good solutions, and fails to find feasible solutions for most of the test problems, two metaheuristics are proposed with different approaches. The first one is combined variable neighborhood search with simulated annealing (VNS-SA) and the second one is hybrid particle swarm optimization (HPSO). Also two lower bounding techniques based on patients' assignment (LB_1) and specialists' assignment (LB_2) are presented. Then the efficiency of the proposed model and algorithms is evaluated using a set of practical-sized test problems. The results showed that VNS-SA is capable of providing high quality solutions in reasonable amount of time for all test problems and outperforms HPSO. Furthermore, the superiority of LB_1 over LB_2 and the lower bound provided by the mathematical model was shown from both the quality and computational time points of view. Finally, some managerial notes and suggestions for extension are presented.

Keywords: Centralized appointment scheduling, distributed echocardiography network, shift scheduling, capacity allocation, variable neighborhood search, particle swarm optimization

1-Introduction

Due to the growth in healthcare expenditures and demand, as wel as patients' expectations of service quality, developing efficient healthcare systems has become very important to governments and healthcare authorities (Hulshof et al., 2012); this gets more critical in outpatient clinics because of the importance of preventive medical measures, shorter hospital length of stays, and providing outpatient services (Cayirli & Veral, 2003).

*Corresponding author

Echocardiography, or simply echo, is the most used cardiac test after electrocardiography (Trang et al., 2019). It uses ultrasound waves to produce moving pictures of the heart. Echo provides prognostic information on the structure and performance of the heart and also facilitates specialists' consultation (Cabell et al., 2004). Furthermore, it is ordered by non-cardiology specialists in many cases. Thus, many parts of a patient's journey depend on echo (Munt et al., 2006). There are several types of echo, which are done by trained sonographers in echo laboratories provided with ultrasound machines, probes, etc.

Echo department is usually a busy place in which patients experience long waiting times; this may lead to delay in the illness recognition, delay in the beginning of cure process, deterioration of the patient condition (Castro & Petrovic, 2012), mental distress (Saure et al., 2012), and patient's dissatisfaction (Pena & Lawrence, 2017). Long waiting times are either caused by insufficient available resources or inefficient patient scheduling (Murray & Berwick, 2003; Saure et al., 2012).

Patient scheduling, or generally appointment scheduling, is a key management tool that can increase resource utilization and decrease patient waiting time. The goal of appointment scheduling is to present an efficient system by optimizing a specific performance measure (Cayirli & Veral, 2003). The most appropriate objective function is a combination of hospitals' revenue management measures and patients' satisfaction measures (Tsai & Teng, 2014). This being the case, we defined our objective function as maximization of profit margin (i.e., revenue of performed echoes minus overtime costs). In order to simplify the calculations, we eliminated fixed costs and just focused on variable costs. In Section 3, we will discuss how this objctive function can improve patients' satisfaction as well.

In an echo department, a group of trained specialists with different specialty and quickness levels work. Also there are echo labs provided with different facilities, where various types of echo with different characteristics in terms of duration, required specialty level and facility are performed. Since an echo requires a lab and a specialist to be performed, we are going to deal with a dual-resource appointment scheduling problem. These characteristics and many other complicating factors of an echo department make the echocardiography appointment scheduling very challenging. A manual and empirical process to develop a schedule for this system is very time-consuming and probably limited to find a feasible solution without focusing on optimality. Therefore, it may lead to long waiting time, waste of capacity, and increase of overtime. Furthermore, it probably does not consider specialists' preferences.

We developed the idea of this paper based on our interviews and observations in Tehran Heart Center (THC). As our motivating case, THC is one of the most advanced and best-equipped diagnostic and therapeutic cardiology centers in Iran and the Middle East. It contains 460 inpatient beds. During the past 16 years, around 2,394,122 outpatients and 296,857 inpatients have received services in THC. Also near 544,933 echocardiography cases of different types have been performed in its echo department. Echo labs of this department are provided with modern imaging equipment and the specialists are from the best in their field ("about Tehran Heart Center," 2017). Based on conducted interviews with the authorities of THC, the echo department is the most crowded department of the hospital. Demands for different types of echo are from outpatients referred to this department from other clinics and hospitals all over the country plus inpatients of THC. Because of the huge demand for echo, inpatients have to experience prolonged length of stay. Also outpatients have to either experience long indirect waiting time (the average interval between requests for appointment by patients to the actual dates of appointment), or leave the department without setting appointments, since there is no available capacity in their requested time. All these cause patients' dissatisfaction and disrupt patients' flows, especially in hospitalization step. In addition, when patients cannot set appointments, they have to search for another reliable center. Finding a center in which the requested echo can be performed with acceptable quality and appropriate insurance is very difficult. It gets more difficult for patients coming from other cities.

Regarding what mentioned above, this idea can be developed that several echo departments/clinics cooperate with each other in terms of a distributed network to generate synergy. The motivation of this study is to present an efficient centralized master schedule for the echocardiography departments/clinics, allowing hospitals/clinics interaction from a system-wide perspective. To keep the quality of this network, these centers should be at the same quality level and reliable to each other.

Centralized management tools are critical in healthcare (Heath, 2017). Based on a case study by Parallon (2013), centralized scheduling has brought both consistency and cost efficiency to a 45-hospital group. Some other benefits provided by centralized scheduling include greater revenue resulting from improved capacity utilization, ability to generate data and measure performance, improved physicians' and patients' satisfaction, expanded access to care, easier addressing billing concerns (Reliasmedia, 2014) and streamlining patients flow by letting them choose the most convenient location and time for their treatment (Scisolutions, 2017).

Our developed centralized appointment scheduling system i) determines the assignment of specialists to different labs in each echo department or clinic in each shift, ii) allocates the available capacity to different groups of inpatients and outpatients, iii) predicts the required overtime, and finally, iv) maximizes profit margin. According to these outputs, we named our proposed schedule *Centeralized Master Echocardiography Schedule* (*CMES*). Decisions addressed in this paper are concerned with specialists' assignment to shifts and labs, and patients' assignments to shifts, labs, and specialists. So, it is required to handle the problems of shift scheduling and capacity allocation, taking the possibility of overtime into account.

The main contributions of this paper are as follows:

- Designing and proposing an efficient master schedule for centralized management of a distributed network of echo departments/clinics
- Joint shift scheduling and capacity allocation in the dual-resource echo environment with non-identical resources of the same type, considering the preferences and possibility of overtime
- Presenting a Mixed Integer Linear Programming (MILP) model along with two metaheuristics and lower bounding techniques, and finding the most efficient ones for CMES.

2-Literature review

Due to the increasing demand and expenditures in healthcare, designing and organizing processes in order to provide high-quality care and optimize available resource utilization has turned into a very important and challenging task for healthcare authorities (Hulshof et al., 2012). Appointment scheduling as a key management tool has attracted much attention, especially during the recent decades. Among the earliest researches on appointment scheduling, the works of Bailey (1952) and Welch and Bailey (1952) can be mentioned. The complicated inherent of appointment scheduling problems in healthcare lies in the fact that many conflicting goals and a variety of constraints regarding efficiency, cost, quality of care, capacity of resources, preferences, overtime, waiting time, etc. should be taken into account.

Decisions made in healthcare contexts with respect to considered goals, constraints and requirements are usually classified into three levels of strategic, tactical and operational; however, there are no clear boundaries for these levels (Aringhieri et al., 2015). Strategic decisions (long-term decisions) deal with designing the general structure of the appointment scheduling system. Tactical decisions (medium-term decisions) address the scheduling and general capacity allocation problems for different groups of patients. Finally, operational decisions (short-term decisions) handle the scheduling and sequencing problems of all patients (Ahmadi-Javid et al., 2017). Decisions addressed in this paper, concerned with specialists' assignment to shifts and labs and patients' assignments to shifts, labs and specialists, can be categorized primarily as tactical.

As mentioned, to develop a CMES, it is required to consider shift scheduling and capacity allocation problems simultaneously. Hence, the first part of this section has been allocated to review the most related papers on shift scheduling and capacity allocation.

As for shift planning and scheduling, usually a lot of regulations and requirements should be considered like providing standard level of patient care, hospitals' policies, labor laws, educational opportunities, and preferences. Hamid et al. (2018) considered nurse scheduling problem. They presented a mathematical model with the objectives of reducing costs, maximizing nurses' satisfaction, and balancing the workload of nurses. They developed a two-step procedure to solve the problem. Rabani and Niyazi (2017) proposed an approach based on graph coloring and bipartite graph concept to solve the shift

scheduling problem of nurses. Volland et al. (2017) combined flexible shift scheduling with a task scheduling problem. They developed a mixed integer model to optimize the number of logistic assistants and a column generation-based approach to find optimal solutions. Hong et al. (2018) presented an integer programming-based approach embedded within a recursive algorithm to provide a set of Pareto-dominant solutions for shift scheduling in an emergency department. Smalley and Keskinocak (2016) presented two integer programming models to construct rotation schedules for resident physicians at academic teaching hospitals. Brech et al. (2019) discussed the problem of constructing monthly training schedules for medical residents with the goal of minimizing the tardiness of training. They developed a mixed integer programming model and a metaheuristic combining benders decomposition and ant colony optimization to find the number and type of surgical procedures a resident performs each month.

The next decision is capacity allocation to patient groups, which is a tactical decision. Tactical decisions mainly deal with increasing productivity and accessibility of high-quality care services (Ahmadi-Javid et al., 2017). Nguyen et al. (2015) discussed allocating the capacity of the healthcare providers to the patient demand in a re-entry system regarding predetermined appointment lead-times for patients. They presented a mixed-integer programming model for planning capacity with the objective of minimizing the maximum required capacity. They solved this model with a network flow approach based on Branch and Cut algorithm. Zhou et al. (2017) discussed allocating the capacity of imaging facilities to different types of patients considering equity. They proposed a nonlinear mixed integer programming model with the objective of maximizing revenue. They defined the problem as a M/D/n queuing system and developed an approximated model. Zhou et al. (2018) focused on allocating the limited capacity of wards to different types of patients with the goal of maximizing both revenue and equity. They used a data-driven discrete-event simulation model to propose a multi-objective integer linear programming model. They also developed an adaptive improved ε-constraint algorithm and a multi-objective genetic algorithm to solve the problem. Bakker and Tsui (2017) presented a data-driven algorithmic approach to allocate specialists to activities and patient groups. They evaluated the performance of their approach using discrete-event simulation. It is worth mentioning that, in the present study, we determined the number of echos of each type that should be performed in each shift and by each specialist while making decisions on capacity allocation. Determining the number of patients that can be scheduled in each consultation session is another tactical decision that has been discussed in some studies. For example, LaGanga and Lawrence (2012) developed a flexible appointment scheduling model to mitigate the negative effects of no-shows. They presented an effective approach to create near-optimal overbooked appointment schedules to balance the advantages of serving additional patients with the potential costs of waiting time and overtime.

Moreover, we reviewed the most relevant literature considering appointment scheduling in operating room (OR) or specialty clinics. Vogl et al. (2019) considered the problem of scheduling radiotherapy treatment appointments with the aim of minimizing the operation time of the particle beam, while simultaneously minimizing penalties of time window violations. They developed three metaheuristic approaches based on genetic algorithm to solve the problem. Zhang et al. (2019) discussed outpatient appointment scheduling when a team of clinicians, technicians, and staff provide treatment in a single patient visit. The objective was minimizing a combination of closing time and total patient waiting time. They developed a two-stage stochastic optimization model to solve the problem. Vali Siar et al. (2017) addressed a multi-period and multi-resource OR scheduling and rescheduling problem with elective and semi-elective patients. The objective was minimization of tardiness, idle-time and overtime. They proposed a scheduling-rescheduling framework based on rolling horizon approach. The core of their proposed framework was a mixed-integer linear programming model that incorporated pre-operative holding unit beds and recovery beds. Klassen and Yoogalingam (2019) considered the addition of mid-level service providers such as physician assistants or nurse practitioners as less costly capacity. They discussed how scheduling policies for single-stage environment could be adjusted for a multi-stage environment, and proposed some scheduling rules using simulation optimization. Nazif (2018) discussed OR surgery scheduling problem. They considered uncertainty in surgery durations by means of fuzzy numbers and proposed an ant colony metaheuristic algorithm for sequencing patients and allocating

resources. Tohidi et al. (2019) presented an integrated physician and clinic scheduling problem in ambulatory cancer treatment polyclinics. They developed a multi-objective optimization model and a hybrid algorithm based on iterated local search and variable neighborhood descent methods to solve the problem. Hamid et al. (2019) addressed the scheduling problem of inpatient surgeries. To improve the compatibility level within the surgical teams in ORs, the decision-making styles of the surgical team members were considered. They developed two metaheuristics based on genetic and particle swarms optimization to find Pareto solutions. Also, as examples of scheduling in a network of hospitals, the works of Santibáñez et al. (2007) and Roshanaei et al. (2017) can be mentioned. Santibáñez et al. (2007) presented a system-wide model to create trade-offs between OR availability, bed capacity, surgeons' booking privileges, and wait lists. They developed a mixed integer programming model to schedule surgical blocks of ORs considering post-surgical capacities. Roshanaei et al. (2017) developed logic-based Benders' decomposition approaches and a cut propagation mechanism to schedule patients and ORs across a network of hospitals. The goal was scheduling the patients with the highest priority scores in the current planning horizon and determining the number of surgical suites and ORs required to create the schedule at minimum cost.

From another standpoint, we reviewed the papers that focused on process improvement regarding echocardiography. Katsi et al. (2014) proposed descriptive productivity measures of echocardiography studies. According to their analysis, the number of studies per physician per day is a good measure to assess productivity. Bakshi (2013) applied a descriptive study and case study method with intensive observation of patient flow, delays, and short comings in patient movement and workflow in a cardiology department. Deploying process reengineering methods, he examined the processes of the existing system and recommended necessary suggestions. Geronimo (2017) focused on improving access to stress echo in an emergency department via observations, time studies and Plan-Do-Study-Act process. Deploying root cause analysis and team discussions, she proposed to add a stress echo lab, purchase new stress test facility, and change the schedule of some nurses. Gandhi (2013) discussed the appointment scheduling problem in an echo department in order to increase the number of scans per day. To deal with this problem, she developed a simulation model and six different policy change scenarios. All scenarios showed significant improvements, specially the scenario which eliminated sonographer schedules.

In conclusion, many remarkable studies have been performed in the field of healthcare appointment scheduling or echocardiography productivity improvement. However, to the best of our knowledge, there is no research on the centralized appointment scheduling in a network of echocardiography departments in a multi-hospital system, considering shift scheduling and capacity allocation problems simultaneously.

3-Problem description and formulation

In this section, our developed MILP for CMES is presented. As already mentioned, we considered a two-part weighted objective function to maximize profit margin. The first part is the maximization of revenue through maximization of the weighted number of performed echos. This part increases the hospital revenue and decreases underutilization of resources. Furthermore, it provides the highest possible level of access to care for patients, and consequently, decreases patients' indirect waiting times and thus increases patients' satisfaction. The second part of our objective function is the minimization of overtime. This part decreases hospitals' variable costs and personnel dissatisfaction.

In this study, we propose a centralized shift scheduling and capacity allocation system for a distributed network of several echo departments and clinics (at the same quality level and reliable to each other) in cooperation. Various types of echos, each requiring specific facilities and specialty level, are performed in these departments/clinics. Each department/clinic is composed of several echo labs. Different facilities are located in each lab; thus, the echo labs are not exactly the same. Moreover, there are several specialists with different specialty and quickness levels in this network. These specialists work at one or more of these centers. They are only capable of performing echos compatible with their specialty level. They work at echo labs in predefined shifts. In this network, demands for different types of echo are either from the

inpatients of this network, or from outpatients. Obviously, echos of inpatients should be performed in the echo department of the hospitals they are in; however, there is no such limitation for outpatients.

In this study, we attempt to assign specialists to different shifts and echo labs, considering their characteristics, preferences, availability, the hospital they work in, and the allowable usual and overtime hours. Also we allocate the available capacity of the network (specialists and echo labs) to different groups of inpatients and outpatients. As mentioned, this study focuses on tactical level decisions and does not directly addresses operational level decisions; however, the structure of the output let us determine the date of appointments, considering patients' or their doctors' preferences, as well. In our CMES, it is clear that what parts of capacity are available for each inpatient/outpatient. Hence, for the demand of each inpatient, his/her doctor decides the most appropriate date and specialist amongst the available options to have his/her echo done. Additionally, each outpatient can choose the most appropriate option amongst the available ones based on his/her preferences. Patient's preference might be having his/her echo done by a specific specialist, in a specific date, in a specific hospital/clinic, or just as soon as possible.

Some other assumptions of our problem are: a) each type of echo can be performed only in a lab with the required facilities, and only by a specialist with the required specialty level, b) at any time and in each echo lab, at most one specialist can perform echo, and at most one patient can be visited, c) each specialist in any shift can be assigned to at most one lab, d) due to difference in the specialists' quickness level, the echo duration not only depends on the echo type, but also on the quickness of the specialist that performs it, e) often two shifts are defined in each day; therefore, the planning horizon of the developed schedule is made up of date-shift combinations. In the rest of the paper, for the sake of simplicity, combinations of date-shift are just referred to as shifts. Furthermore, we allow the possibility of overtime on different shifts of the planning horizon.

Many hospitals repeat their weekly cyclic schedule over a month or even a period of several months. Also many hospitals maintain it with minor modifications until a substantial change occurs (Penn et al., 2017). For our proposed CMES, the planning horizon of one week seems to be apporopriate. Since the weekly demand in each month does not differ significantly from one week to another, an efficient master schedule can be obtained by our proposed method for one week; this schedule can be repeated for a month or until demand rates or any other essential parameters of the system change.

Since the decisions made at the tactical level are less affected by the uncertainty in patients' arrival and service times (Ahmadi-Javid et al., 2017), we presented a deterministic mixed integer linear programming model for the research problem. Sets, parameters and variables of the model are as follows:

Sets and indices

T	Set of all shifts in the planning horizon, indexed by t	$T = \{1,2,\dots,t,\dots,\lvert T\rvert\}$
J	Set of all specialists, indexed by j	$J = \{1,2,\dots,j,\dots,\lvert J\rvert\}$
I	Set of all patient types (echo types), indexed by i	$I = \{1,2,\dots,i,\dots,\lvert I\rvert\}$
H	Set of all hospitals (or clinics), indexed by h	$H = \{1,2,\dots,h,\dots,\lvert H\rvert\}$
B	Set of all situations for inpatients and outpatients, indexed by b $b = h$ for inpatients in hospital h $b = \lvert H\rvert + 1$ for outpatients	$B = \{1,2,\dots,b,\dots,\lvert H\rvert + 1\}$
L	Set of all echo labs, indexed by l	$L = \{1,2,\dots,l,\dots,\lvert L\rvert\}$

Parameters

S_{ij}	$\begin{cases} 1, \text{if echo type } i \text{ can be performed by specialist } j, \text{according to the specialist's specialty level} \\ 0, \text{otherwise} \end{cases}$
E_{il}	$\begin{cases} 1, \text{if the facility needed by echo type } i \text{ is available in echo lab } l \\ 0, \text{otherwise} \end{cases}$
N_{lb}	$\begin{cases} 1, \text{if performing echo of a patient in situation } b \text{ in echo lab } l \text{ is possible} \\ 0, \text{otherwise} \end{cases}$
P_{ij}	Expected duration of echo of type i performed by specialist j
WH_{l_i}	Regular available time of echo lab l in shift t

MO_j Total maximum allowable overtime for specialist j in the planning horizon

TO_{lt} Maximum allowable overtime of echo lab l in shift t

D_{ib} Total demand of patients of type i in situation b in the planning horizon

k_{ib} The minimum percentage of echos of type i and situation b that should necessarily be performed within this planning horizon

CO Overtime cost per hour

W_{ib} Obtained revenue by performing each echo of type i in situation b

ZL_j Minimum allowable hours to work for specialist j in the planning horizon

ZU_j Maximum allowable hours to work for specialist j in the planning horizon

O_{jh} $\begin{cases} 1, \text{if specialist } j \text{ works at hospital } h \\ 0, \text{otherwise} \end{cases}$

F_{jth} $\begin{cases} 0.1 \leq F_{jth} < 1, \text{ the preferance of specialist } j \text{ to work at hospital } h \text{ in shift } t \\ 0, \text{if specialist } j \text{ is strictly reluctant to work at hospital } h \text{ in shift } t \end{cases}$

M A large number

Decision variables

Y_{jtl} $\begin{cases} 1, \text{if specialist } j \text{ is assigned to lab } l \text{ in shift } t \\ 0, \text{otherwise} \end{cases}$

X_{iblt} Number of echoes of type i in situation b, assigned to shift t to be performed in lab l

OH_{tj} Length of overtime of specialist j in shift t

Note that $N_{lb} = 1$ if either $b = |H| + 1$ (i.e., for outpatients) or l is a lab of hospital $h = b$ for inpatients in hospital $h = b$. The $MILP$ is developed as follows:

$$\max \; Z = \sum_{i \in I, b \in B, t \in T, l \in L} (W_{ib} X_{iblt}) - CO \sum_{t \in T, j \in J} OH_{tj} \tag{1}$$

$$\sum_{i \in I, b \in B} X_{iblt} P_{ij} \leq WH_{lt} + OH_{tj} + M(1 - Y_{jlt}) \qquad \forall t \in T, j \in J, l \in L \tag{2}$$

$$\sum_{t \in T} OH_{tj} \leq MO_j \qquad \forall j \in J \tag{3}$$

$$OH_{tj} \leq \sum_{l \in L} TO_{lt} Y_{jlt} \qquad \forall t \in T, j \in J \tag{4}$$

$$\sum_{t \in T, l \in L} X_{iblt} \geq k_{ib} D_{ib} \qquad \forall i \in I, b \in B \tag{5}$$

$$\sum_{t \in T, l \in L} X_{iblt} \leq D_{ib} \qquad \forall i \in I, b \in B \tag{6}$$

$$X_{iblt} \leq D_{ib} \sum_{j \in J} S_{ij} Y_{jlt} \qquad \forall i \in I, b \in B, l \in L, t \in T \tag{7}$$

$$\sum_{t \in T} X_{iblt} \leq D_{ib} E_{il} N_{lb} \qquad \forall i \in I, b \in B, l \in L \tag{8}$$

$$\sum_{j \in J} Y_{jlt} \leq 1 \qquad \forall t \in T, l \in L \tag{9}$$

$$\sum_{l \in L} Y_{jlt} \leq 1 \qquad \forall t \in T, j \in J \tag{10}$$

$$\sum_{t\in T, l\in L} WH_{lt}\mathrm{Y}_{jlt} \leq \mathrm{ZU}_j \qquad \forall j \in J \qquad (11)$$

$$\sum_{t\in T, l\in L} WH_{lt}\mathrm{Y}_{jlt} \geq \mathrm{ZL}_j \qquad \forall j \in J \qquad (12)$$

$$\sum_{t\in T} \mathrm{Y}_{jlt} \leq T \sum_{h\in H} O_{jh}\, N_{lh} \qquad \forall j \in J, l \in L \qquad (13)$$

$$\mathrm{Y}_{jlt} \leq M \sum_{h\in H} F_{jth}\, N_{lh} \qquad \forall t \in T, j \in J, l \in L \qquad (14)$$

$$Y_{jlt} = 0 \; or \; 1 \qquad \forall t \in T, j \in J, l \in L \qquad (15)$$

$$X_{iblt} \geq 0, int \qquad \forall i \in I, t \in T, l \in L, b \in B$$

$$OH_{tj} \geq 0 \qquad \forall t \in T, j \in J$$

The objective function, as presented in equation (1), is defined to maximize the profit margin. The first part calculates the obtained revenue, and the second part determines the overtime cost. Constraint set (2) ensures that the total time of all echoes assigned to each shift-lab is less than or equal to the sum of regular shift's length and the overtime of the specialist assigned to that lab in that shift. Constraint sets (3) and (4) determine the overtime of specialist *j* in shift *t*. Constraint set (3) ensures that the total overtime of specialist *j* does not exceed the maximum allowable overtime for that specialist. Constraint set (4) guarantees that if specialist *j* does not work at shift *t*, then the decision variable related to the overtime of specialist *j* in shift *t* should be equal to zero. In addition, it accounts for the limitation on the maximum allowable overtime in shift *t* and lab l. Constraint set (5) is incorporated into the model to make sure that the total number of scheduled patients of each echo type and situation covers at least a predefined percentage of that echo type and situation in the planning horizon. Constraint set (6) accounts for the limitation on the total demand of each echo type. Constraint set (7) supports the fact that the assignment of patients of echo type *i* to specialist *j* is possible, if and only if specialist *j* has the required specialty level related to echo type *i*. Constraint set (8) guarantees that the assignment of patients of echo type *i* to lab *l* is possible, if and only if lab *l* is equipped with the required facilities. Besides, it assures the fact that echoes of inpatients in hospital *b* can only be performed in the labs (with required facilities) of hospital *b*, while echoes of outpatients can be assigned to any lab (with required facilities). Constraint set (9) ensures that at most one specialist should be assigned to each shift-lab combination. Constraint set (10) supports the fact that each specialist in any shift can be assigned to at most one lab. Constraint sets (11) and (12) guarantee that the sum of regular hours that specialist *j* works is in the range of minimum and maximum allowable hours for that specialist in the planning horizon. Constraint set (13) assures that a specialist can only be assigned to the labs of hospitals that he/she works at. Constraint set (14) is incorporated into the model to make the assignment of a specialist to a specific lab and shift impossible if he/she is strictly unwilling to work at that hospital in that shift. Finally, Constraint set (15) defines the binary, integer and continuous variables.

4-Solution methods

Assignment of labs and shifts to specialists is similar to the Knapsack Problem since there are several specialists (multiple knapsacks) with limitation on the total number of hours to work (capacity of each knapsack). Also there are several lab-shift combinations with different lengths (items with different sizes) that should be assigned to each specialist (be put into each knapsack). As a specialist cannot work at two labs in one shift, lab-shift combinations with the same shift cannot be assigned to the same specialist. This is similar to the Knapsack Problem with incompatible items in which incompatible items cannot be put in the same knapsack. The Knapsack Problem with incompatible items is called Knapsack Problem with Conflict Graph (KPCG), which is an extension of basic Np-Hard Knapsack problem (Bettinelli et al., 2017). So, the problem discussed in this study is Np-Hard.

This being the case, we focused on developing efficient algorithms to heuristically solve the problem. Such algorithms are supposed to accomplish three tasks: 1) specialists' assignment to lab-shift combinations, 2) patients' assignment to lab-shift-specialist combinations, and 3) determination of required overtime for each specialist in each shift. In the following subsections, our proposed VNS-SA and HPSO, which are developed with different approaches, are described in detail.

4-1-Combined Variable Neighborhood Search and Simulated Annealing algorithm (VNS-SA)

Variable Neighborhood Search (VNS) algorithm is a metaheuristic proposed by Mladenović and Hansen (1997). VNS selects a set of neighborhood structures and an initial solution. Through a shaking procedure, it systematically changes the neighborhood structure in order to escape from getting trapped in the local optima, and by a local search procedure, it searches the neighborhood. Simulated annealing (SA) is also a probabilistic method proposed by Kirkpatrick et al. (1983). SA starts the search procedure from an initial solution and creates a neighbor solution by implementing the neighborhood search structure on the current solution. It accepts improving solutions and a fraction of non-improving solutions with the hope of escaping the local optima. We decided to combine VNS and SA such that VNS is responsible for selecting the neighborhood structure and SA conducts the local search. Furthermore, when the main phase of VNS-SA stops, an improvement phase is implemented.

In the proposed VNS-SA, at each stage of the algorithm, a solution determining the assignment of specialists to labs and shifts is produced. Thereafter, the fitness function calculates the best obtainable profit margin for the recently produced specialists' assignment to lab-shift combinations. Actually, this fitness function solves a mathematical model to find the optimal (or near optimal) assignment of patients to lab-shift-specialist combinations and required overtime. In other words, each solution has three parts. The first part, which is created, updated and determined in the body of VNS-SA, is specialists' assignment to lab-shift combinations. The second part is patients' assignment to lab-shift-specialist combinations, and the third part is required overtime for each specialist in each shift. Both the second and third parts of the solution are calculated and determined by the fitness function of VNS-SA. In the rest of this section, and for the sake of simplicity, specialists' assignment to lab-shift combinations (i.e., the first part of solution) is called "solution". Also all the first, second and third parts together, will be referred to as a "comprehensive solution". A flowchart for the proposed algorithm is provided in Appendix A.

4-1-1- Solution representation

To represent each solution, a table of T rows and J columns is applied. The cell in row t and column j displays the lab to which specialist j is assigned, in shift t. Figure 1 demonstrates an example with four shifts, three labs and six specialists. A cell value equal to zero implies that the corresponding specialist of that column is not assigned to any lab in the corresponding shift of that row.

		Specialists					
		$\boldsymbol{j=1}$	$\boldsymbol{j=2}$	$\boldsymbol{j=3}$	$\boldsymbol{j=4}$	$\boldsymbol{j=5}$	$\boldsymbol{j=6}$
Shifts	$\boldsymbol{t=1}$	1	0	2	3	0	0
	$\boldsymbol{t=2}$	3	1	0	2	0	0
	$\boldsymbol{t=3}$	0	0	1	0	2	3
	$\boldsymbol{t=4}$	0	2	0	0	3	1

Fig 1. An example for solution representation in VNS-SA

4-1-2- Initial solution construction

The initial solution must be a feasible assignment of specialists to labs and shifts such that the following conditions are met:

1. At most one specialist should be assigned to each shift-lab combination.
2. Each specialist, in any shift, can be assigned to at most one lab.
3. Each specialist can only be assigned to the labs of hospitals that he/she works at.
4. The assignment of a specialist to a lab and shift is impossible if he/she is strictly unwilling to work at that hospital in that shift.
5. The sum of regular hours that each specialist works should be in the predefined range of minimum and maximum allowable hours for that specialist.

Therefore, a random solution is likely infeasible. In our VNS-SA, a mathematical model (16-24) is applied to construct a feasible initial solution. ε_{lt} is defined as a binary variable such that $\varepsilon_{lt} = 1$ if lab l is empty at shift t, otherwise $\varepsilon_{lt} = 0$. In this model, Constraint set (19) guarantees the first condition mentioned above. Constraint sets (20), (21) and (22) guarantee the second, third and fourth conditions, respectively. Finally, Constraint sets (23-24) guarantee the fifth condition.

Taking the above constraints into consideration does not suffice to assure a good profit margin, or even a feasible comprehensive solution. An inappropriate specialists' assignment to shifts and labs (though feasible as the first part of solution) may result in infeasibility in determining the second and third parts of the comprehensive solution. Inappropriate specialists' assignment, leading to waste of capacity, can obviously decrease profit margin. Also it might fail to handle the constraint that at least a predefined percentage of echos of each type and situation should necessarily be covered within this planning horizon (constraint set (5)). Hence, we defined three objective functions to increase the quality of the initial solution and to decrease the probability of facing infeasibility for the remaining parts of the comprehensive solution. Using these objective functions leads to better schedule for specialists' assignment and thus better capacity utilization.

The first objective function, as presented in equation (16), minimizes the total number of empty labs. The second objective function (equation (17)) maximizes compatibility between the specialty level of a specialist and the lab he/she is assigned to. Parameter VAL_{jl} is defined as the compatibility between the specialty level of specialist j and the facilities provided in lab l. We defined VAL_{jl} in the range of [1,4] such that higher values of VAL_{jl} mean higher compatibility. Eventually, the latter objective function (equation (18)), maximizes average obtainable revenue. Parameter GW_{jlt} is defined as the estimated average obtainable revenue by assigning specialist j to lab l in shift t. The value of this parameter can be easily and approximately calculated considering the regular time of lab l in shift t, types of echo that can be performed according to the specialty level of specialist j and equipment level of lab l, and quickness level of specialist j.

$$minimize\ F_1 = \sum_{t\in T, l\in L} \varepsilon_{lt} \tag{16}$$

$$maximize\ F_2 = \sum_{t\in T, l\in L, j\in J} VAL_{jl} \mathrm{Y}_{jlt} \tag{17}$$

$$maximize\ F_3 = \sum_{t\in T, l\in L, j\in J} GW_{jlt} \mathrm{Y}_{jlt} \tag{18}$$

$$\sum_{j\in J} \mathrm{Y}_{jlt} + \varepsilon_{lt} = 1 \qquad \forall t \in T, l \in L \tag{19}$$

$$\sum_{l\in L} \mathrm{Y}_{jlt} \leq 1 \qquad \forall t \in T, j \in J \tag{20}$$

$$\sum_{t\in T} \mathrm{Y}_{jlt} \leq T \sum_{h\in H} O_{jh} N_{lh} \qquad \forall j \in J, l \in L \tag{21}$$

$$\mathrm{Y}_{jlt} \leq M \sum_{h\in H} F_{jth} N_{lh} \qquad \forall t \in T, j \in J, l \in L \tag{22}$$

$$\sum_{l\in L,t\in T} WH_{lt}\mathrm{Y}_{jlt} \le ZU_j \qquad \forall j \in J \qquad (23)$$

$$\sum_{l\in L,t\in T} WH_{lt}\mathrm{Y}_{jlt} \ge \mathrm{ZL}_j \qquad \forall j \in J \qquad (24)$$

$$Y_{jlt} = 0 \ or \ 1 \qquad \forall t \in T, j \in J, l \in L \qquad (25)$$

$$\varepsilon_{lt} = 0 \ or \ 1 \qquad \forall t \in T, l \in L$$

To incorporate the three defined objective functions, we used the weighted objective function approach and thus substituted Eq. (16-18) with Eq. (26).

$$minimize\ F = weight_1 \sum_{t\in T,l\in L} \varepsilon_{lt} - weight_2 \sum_{t\in T,l\in L,j\in J} VAL_{jl}\mathrm{Y}_{jlt} - weight_3 \sum_{t\in T,l\in L,j\in J} GW_{jlt}\mathrm{Y}_{jlt} \qquad (26)$$

The experience of solving a large number of test problems verified the necessity of these three objective functions and the significant effect of their weights on the feasibility and quality of solutions.

4-1-3- Fitness function: an algorithm that determines patients' assignment and overtime, and calculates the objective function

At each stage of VNS-SA, whenever a new solution (i.e., specialists' assignment) is generated, a fitness function is called to calculate the objective function. In order to calculate the best obtainable objective function for the current solution, first the most appropriate patients' assignment considering overtime possibility needs to be determined. The following mathematical model (27-35) is proposed to do so. Parameter A_{jlt} demonstrates the current specialists' assignment. $\mathrm{A}_{jlt} = 1$ if specialist j is assigned to lab l in shift t based on recently generated solution, otherwise $\mathrm{A}_{jlt} = 0$. Since the objective function and constraints of this model are similar to equation (1) and Constraint sets (2-8), we avoid explaining them again.

$$\max \sum_{i\in I,b\in B,t\in T,j\in J} W_{ib}X_{iblt} - CO \sum_{t\in T,j\in J} OH_{tj} \qquad (27)$$

$$\sum_{i\in I,b\in B} (X_{iblt} \sum_{\forall j \in J} \mathrm{A}_{jlt}\, P_{ij}) \le WH_{lt} + \sum_{\forall j \in J} \mathrm{A}_{jlt}\, OH_{tj} \qquad \forall t \in T, l \in L \qquad (28)$$

$$\sum_{t\in T} OH_{tj} \le MO_j \qquad \forall j \in J \qquad (29)$$

$$OH_{tj} \le \sum_{l\in L} TO_{lt}\mathrm{A}_{jlt} \qquad \forall t \in T, j \in J \qquad (30)$$

$$\sum_{t\in T,l\in L} X_{iblt} \ge k_{ib}D_{ib} \qquad \forall i \in I, b \in B \qquad (31)$$

$$\sum_{t\in T,l\in L} X_{iblt} \le D_{ib} \qquad \forall i \in I, b \in B \qquad (32)$$

$$X_{iblt} \le D_{ib} \sum_{j\in J} S_{ij}\mathrm{A}_{jlt} \qquad \forall i \in I, b \in B, l \in L, \forall t \in T \qquad (33)$$

$$\sum_{t\in T} X_{iblt} \le D_{ib}E_{il}N_{lb} \qquad \forall i \in I, b \in B, l \in L \qquad (34)$$

$$X_{iblt} \ge 0, int \qquad \forall i \in I, b \in B, t \in T, j \in J \qquad (35)$$

$$OH_{tj} \ge 0 \qquad \forall t \in T, j \in J$$

4-1-4- Neighborhood search structures

The ability of neighborhood structures in searching the feasible solutions area has a key impact on the probability of discovering good solutions. In the proposed VNS-SA, not only should we consider different kinds of constraints while moving from a feasible solution to another, but also we have to take care of the solution in order not to face infeasibility in determining patients' assignment and calculating objective function. This complexity being the case, we applied 10 neighborhood structures to provide more chance for the algorithm to search the feasible solutions area well. Note that the solution area is the discrete area of specialists' assignment to lab-shift combinations.

The proposed neighborhood structures are described as follows. All of them are applied on the specialists' assignment table (explained in Section 4-1-1). These structures are illustrated in Appendix B.

1. Select two specialists j_1 and j_2 randomly. In a randomly chosen shift t, exchange their corresponding cell values (i.e., their assigned labs in shift t) in the specialists' assignment table.
2. Select two specialists j_1 and j_2 randomly such that they are at the same specialty level. In a randomly chosen shift t, exchange their corresponding cell values in the specialists' assignment table.
3. Select three specialists j_1, j_2 and j_3 randomly. In a randomly chosen shift t, exchange their corresponding cell values in the specialists' assignment table.
4. Select two specialists j_1 and j_2 randomly. In each of the two randomly chosen shifts t_1 and t_2, exchange the specialists' corresponding cell values in the specialists' assignment table.
5. Select two specialists j_1 and j_2 randomly such that they are at the same specialty level. In each of the two randomly chosen shifts t_1 and t_2, exchange the specialists' corresponding cell values in the specialists' assignment table.
6. Select two specialists j_1 and j_2 randomly. In each of the three randomly chosen shifts t_1, t_2 and t_3, exchange the specialists' corresponding cell values in the specialists' assignment table.
7. In a randomly chosen sift t, select a part of the table randomly and flip it.
8. In a randomly chosen sift t, select a part of the table randomly and just flip the cells corresponding to the highest specialized specialists.
9. Select two specialists j_1 and j_2 randomly, and exchange their corresponding cell values in all shifts.
10. Select two specialists j_1 and j_2 randomly such that they are at the same specialty level. Exchange their corresponding cell values in all shifts.

Since applying repair mechanisms for the infeasible solutions of this problem is very difficult, the neighborhood structures were defined in a way that the probability of generating an infeasible solution decreases. However, it still might happen. Therefore, we imposed two conditions to be checked whenever a neighborhood structure is implemented. The solutions that did not satisfy these conditions were eliminated.

- Condition 1: The number of empty labs should not increase. In other words, when a new value l' is inserted at the intersection of row t and column j, specialist j should be able to perform echo in lab l' in shift t (according to parameters F_{jth} and O_{jh}).
- Condition 2: The sum of regular hours each specialist works should be in the range of minimum and maximum allowable hours defined for him/her in advance.

4-1-5- Improvement phase

In the proposed algorithm, when the main phase of VNS-SA stops (the stopping criterion is discussed in 4-1-6), an improvement phase is implemented. Table 1 displays several rules that potentially improve the objective function. In each iteration of the improvement phase, one shift t and two specialists j_1 and j_2 are randomly chosen. Depending on the existing feature, the related potentially improving rule(s) are applied (if the generated solution meets the conditions mentioned in section 4.1.4). In case of more than one rule, the solution with better objective function is kept. The newly created solution is accepted only if it has a better objective function value.

Table 1. Potentially improving rules for the improvement phase

	Feature (for the chosen shift)	Potentially improving rule(s)
1	j_1 and j_2 are at the same specialty level assigned to different labs	The lab with longer available time should be assigned to the specialist with higher average speed
2	j_1 and j_2 are at different specialty levels assigned to different labs	The lab with higher equipment level is assigned to the higher specialized specialist
3	j_1 and j_2 are at the same specialty level; one is assigned to a lab and the other one is not assigned to any lab	The lab should be assigned to the specialist with higher average speed
4	j_1 and j_2 are at different specialty levels; one is assigned to a lab and the other is not assigned to any lab	The lab should be assigned to the specialist with higher average speed / The lab should be assigned to the higher specialized specialist

4-1-6- Stopping criterion

In this algorithm, time limitation is chosen as the stopping criterion. A specified percentage of this amount of time $(1-\beta)$ is allocated to the main phase of VNS-SA, and the rest (β) is allocated to the improvement phase. β is tuned in section 5.2.

4-2-Hybrid Particle Swarm Optimization algorithm (HPSO)

In the VNS-SA algorithm, which was described in the previous section, many neighborhood structures were defined to explore the feasible solution area of specialists' assignment more efficiently. Unfortunately, they are probable to cause infeasibility both in specialists' assignment and patients' assignment, when implemented. For this reason, we developed HPSO with the heuristic idea of creating priority tables instead of specialists' assignment tables. In our HPSO, we search within the continuous area of priority tables with the hope of eliminating the need for complicating neighborhood structures and getting rid of infeasible specialists' assignment. More details are provided in Section 4.2.1.

PSO is a population-based algorithm, which was proposed by Kennedy and Eberhart (1995). In this algorithm, each particle is represented by a vector in a multi-dimensional solution space. During the search process, these particles move around the solution space based on their own best experience (*pbest*) and the other particles' best experience (*gbest*). The velocity ($V_i(k)$) and position ($X_i(k)$) of particle i in k^{th} iteration are calculated by the following equations, respectively.

$$V_i(k+1) = wV_i(k) + c_1 r_1\big(pbest_i - X_i(k)\big) + c_2 r_2\big(gbest - X_i(k)\big) \tag{36}$$
$$X_i(k+1) = X_i(k) + V_i(k+1) \tag{37}$$

Where, w, called inertial weight, states how the velocity of each iteration affects the velocity of the next iteration, and c_1 and c_2 are the cognition and social learning factors, respectively. In most researches, the values assigned to c_1 and c_2 are numbers close to 2. r_1 and r_2 are uniformly distributed random numbers in [0,1]. $pbest_i$ is the vector for the best known position of particle i, and $gbest$ is the best position vector of all particles in the whole population.

In our HPSO, similar to VNS-SA, solutions determine the specialists' assignment to labs and shifts. In each iteration of HPSO, several new solutions are produced. Then a fitness function calculates the objective function of each particle. In fact, this function determines patients' assignment and required overtime and calculates the best obtainable profit margin for the recently produced specialists' assignment. Furthermore, a local search mechanism is incorporated into HPSO to improve the performance of the algorithm. The flowchart of HPSO is provided in Appendix C.

4-2-1- Encoding of particles

In the proposed HPSO, the particles that are created and then updated at each stage of the algorithm are priority tables. A priority table determines the priority of each specialist in each shift for choosing a lab or not performing echo in that shift.

Since PSO is applied for solving continuous problems, our particles are continuous priority matrices. Then the Smallest Position Value (SPV) technique is applied to transform a table in continuous space to the one in discrete space. SPV finds each specialist-shift's priority by finding non-decreasing order of the continuous priority table elements. An example is displayed in figure 2. For instance, according to table b of Figure 2, specialist 3 has the first priority to choose his/her lab for shift 3. In the rest of this section, a solution refers to a continuous priority table (similar to table a of figure 2).

		$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
Shifts	$t=1$	9.25	7.12	5.91	8.35	7.41
	$t=2$	2.02	3.75	2.35	5.30	8.70
	$t=3$	6.44	9.33	0.31	2.01	3.51

a

SPV →

		$j=1$	$j=2$	$j=3$	$j=4$	$j=5$
Shifts	$t=1$	14	10	8	12	11
	$t=2$	3	6	4	7	13
	$t=3$	9	15	1	2	5

b

Fig 2. Particle encoding in HPSO

4-2-2- Initial swarm construction

In order to construct the initial swarm, several random continuous priority tables are generated.

4-2-3- Creating a specialists' assignment table out of a priority table

Here, a heuristic algorithm composed of the following phases is proposed to create a specialists' assignment table out of a priority table. This algorithm also strives to assign specialists such that the probability of facing infeasibility in patients' assignment decreases.

Phase 0 (Preliminaries): In each shift, categorize labs in separate sets based on their equipment levels. Then, sort the labs of each set in decreasing order of their regular available time in that shift. Next, create a list of specialist-shift combinations that should be managed based on the priority table.

Phase 1 (Ensuring the minimum allowable working hours for specialists): Consider the list of specialist-shift combinations from the beginning. For each specialist-shift combination, if the minimum allowable working hours of that specialist is satisfied, pass this specialist-shift and go to the next one; otherwise, assign a lab to that specialist in that shift (according to the *assignment phase*). Continue to the end of the list or until the minimum allowable hours for all specialists are satisfied.

Phase 2 (Completing the solution): Consider the list of specialist-shift combinations from the beginning. For each specialist-shift combination, assign a lab to that specialist in that shift (according to the *assignment phase*). Continue to the end of the list or until all labs in all shifts are assigned.

Assignment phase (For specialist j' in shift t'): In shift t', sort the sets of labs based on decreasing compatibility with specialist j' (compatibility i.e., VAL_{jl} is explained in Section 4.2.1). Assign the first lab of this list to specialist j' in shift t' such that specialist j' can work at the corresponding hospital, specialist j' is not reluctant to work at that hospital in that shift, and finally, maximum allowable working hours of specialist j' is not exceeded. Remove that lab from the list of available labs in shift t' and remove this specialist-shift from the list of specialist-shift combinations that should be managed. If no lab is assigned, just remove this specialist-shift from the list of specialist-shift combinations that should be managed.

4-2-4-Fitness function: an algorithm that determines patients' assignment and overtime, and calculates objective function

For calculating the objective function of each particle, first the priority table should be turned into a specialists' assignment table by implementing the mechanism explained in section 4.2.3. Then the

function explained in section 4.1.3 is called to determine patients' assignment and required overtime, and finally, calculate the objective function.

4-2-5-Updating particles

According to Poli et al. (2007), a multiplier x imposed to equation (36) can accelerate the convergence of PSO. If the condition represented by equation (38) is met, x is calculated by equation (39). Consequently, equation (40) is applied to calculate the new velocity. Next, the positions of particles are updated by equation (37).

$$C = c_1 + c_2 > 4 \tag{38}$$

$$x = \frac{2}{C - 2 + \sqrt{C^2 - 4C}} \tag{39}$$

$$V_i(k+1) = x\left(wV_i(k) + c_1 r_1\left(pbest_i(k) - X_i(k)\right) + c_2 r_2\left(gbest(k) - X_i(k)\right)\right) \tag{40}$$

4-2-6-Local search mechanism

The preliminary experiments indicate that a basic PSO may be trapped in the local optima during the search into the solution space (Poli et al., 2007). Incorporation of a local search scheme into PSO reduces the probability of converging the solutions into the local optima. In each iteration of the proposed HPSO, a local search is performed on all particles by the following random and intelligent neighborhood structures. If a better solution is found in the neighborhood of a particle, the particle is substituted by this new solution.

Random neighborhood structures:

1- Choose two columns of priority table randomly, and then change their values by crossover operation.
2- Choose two rows of priority table randomly, and then change their values by crossover operation.

Intelligent neighborhood structures:

3- Choose two specialists such that they are at the same specialty level. Exchange their priorities in any shift in which the specialist with higher average speed has lower priority.
4- Choose two specialists such that they are at different specialty levels. Exchange their priorities in any shift in which the specialist with higher specialty level has lower priority.

4-2-7-Stopping criterion

Similar to the previous algorithm, time limitation is chosen as the stopping criterion in HPSO.

5-Lower bounding techniques

In this section, two lower bounding techniques are proposed for the mentioned problem.

5-1- LB_1 technique; lower bounding technique regarding patients' assignment constraints

Considering the demand of echo types and total available time of specialists and labs, this technique estimates the highest total obtainable profit margin. In this technique, for different sets of echo types, two constraints should be considered: first, the total time of all performed echoes of this set is less than or equal to the sum of available times of all echo labs provided with the required facilities. Second, the total time of all performed echoes of this set is less than or equal to the sum of available times of all specialists with required specialty level. These sets of constraints should be considered for both echoes performed in regular time and echoes performed in overtime. To make it clear, the proposed model is developed for a simple example. Assume that U_{ib} and V_{ib} are the number of echoes of type i in situation b performed in regular time and overtime, respectively. To solve the model easier, we assumed these parameters to be

continues. $minP_{ib}$ is the shortest time that an echo of type i in situation b can be performed. Other parameters of the example are provided in table 2.

Table 2. Parameters of the example for LB_1

Echo types	Labs				Specialists			
	Types	Equipment for echo types	Total regular available time for each b	Total available overtime for each b	Types	Specialty for echo type	Total regular available time for each b	Total available overtime for each b
$I = \{1,2,3\}$	1	$\{1,2\}$	LW_{1b}	LO_{1b}	1	$\{1\}$	SW_{1b}	SO_{1b}
	2	$\{1,3\}$	LW_{2b}	LO_{2b}	2	$\{1,2,3\}$	SW_{2b}	SO_{2b}

The model is developed as follows. Equation (41) maximizes the profit margin regarding overtime cost. Constraint set (42) accounts for the limitation on total demand of each type in each situation. Constraint sets (43-45) guarantee that the total times of all performed echoes of the specified sets in regular time are less than or equal to the sum of regular available times of all echo labs provided with the required facilities. Constraint sets (46-48) assure the same rule about the echoes performed in overtime. Constraint sets (49-50) guarantee that the total times of all performed echoes of the specified sets in regular time are less than or equal to the sum of regular available times of all specialists with required specialty level. Constraint sets (51-52) assure the same rule about the echoes performed in overtime.

$$maximize\ LB_1 = \sum_{b=1}^{B}\sum_{i=1}^{I}(W_{ib}U_{ib} + (W_{ib} - CO.minP_{ib})V_{ib}) \tag{41}$$

$$U_{ib} + V_{ib} \leq D_{ib} \qquad \forall i \in I, b \in B \tag{42}$$

$$\sum_{i\in\{1,2,3\}} U_{ib}minP_{ib} \leq LW_{1b} + LW_{2b} \qquad \forall b \in B \tag{43}$$

$$\sum_{i\in\{2\}} U_{ib}minP_{ib} \leq LW_{1b} \qquad \forall b \in B \tag{44}$$

$$\sum_{i\in\{3\}} U_{ib}minP_{ib} \leq LW_{2b} \qquad \forall b \in B \tag{45}$$

$$\sum_{i\in\{1,2,3\}} V_{ib}minP_{ib} \leq LO_{1b} + LO_{2b} \qquad \forall b \in B \tag{46}$$

$$\sum_{i\in\{2\}} V_{ib}minP_{ib} \leq LO_{1b} \qquad \forall b \in B \tag{47}$$

$$\sum_{i\in\{3\}} V_{ib}minP_{ib} \leq LO_{2b} \qquad \forall b \in B \tag{48}$$

$$\sum_{i\in\{1,2,3\}} U_{ib}minP_{ib} \leq SW_{1b} + SW_{2b} \qquad \forall b \in B \tag{49}$$

$$\sum_{i\in\{2,3\}} U_{ib}minP_{ib} \leq SW_{2b} \qquad \forall b \in B \tag{50}$$

$$\sum_{i\in\{1,2,3\}} V_{ib}minP_{ib} \leq SO_{1b} + SO_{2b} \qquad \forall b \in B \tag{51}$$

$$\sum_{i\in\{2,3\}} V_{ib} minP_{ib} \leq SO_{2b} \qquad \forall b \in B \qquad (52)$$

$$U_{ib} \geq 0 \qquad \forall i \in I, b \in B$$
$$V_{ib} \geq 0 \qquad \forall i \in I, b \in B \qquad (53)$$

5-2- LB_2 technique; lower bounding technique regarding specialists' assignment constraints

In this technique, first, the highest obtainable revenue of assigning specialist j to shift t and lab l, named Wmax_{jlt}, should be calculated according to the available time of lab l in shift t, as well as the specialty and quickness level of specialist j. Then the following model is solved to yield a lower bound for our main problem. Equation (54) maximizes the approximate revenue. Constraint sets (55-60) are similar to (9-14), so they are not explained again here.

$$maximize\ LB_2 = \sum_{j\in J, l\in L, t\in T} \mathrm{Wmax}_{jlt} \mathrm{Y}_{jlt} \qquad (54)$$

$$\sum_{j\in J} \mathrm{Y}_{jlt} \leq 1 \qquad \forall t \in T, l \in L \qquad (55)$$

$$\sum_{l\in L} \mathrm{Y}_{jlt} \leq 1 \qquad \forall t \in T, j \in J \qquad (56)$$

$$\sum_{l\in L, t\in T} WH_{lt} \mathrm{Y}_{jlt} \leq ZU_j \qquad \forall j \in J \qquad (57)$$

$$\sum_{l\in L, t\in T} WH_{lt} \mathrm{Y}_{jlt} \geq ZL_j \qquad \forall j \in J \qquad (58)$$

$$\sum_{t\in T} \mathrm{Y}_{jlt} \leq T \sum_{h\in H} O_{jh} N_{lh} \qquad \forall j \in J, l \in L \qquad (59)$$

$$\mathrm{Y}_{jlt} \leq M \sum_{h\in H} F_{jth} N_{lh} \qquad \forall t \in T, j \in J, l \in L \qquad (60)$$

$$Y_{jlt} = 0\ or\ 1 \qquad \forall t \in T, j \in J, l \in L \qquad (61)$$

6-Computational experiments

In this section, the performance of the developed MILP model, metaheuristic algorithms and lower bounding techniques is evaluated and compared. MILP model was coded in CPLEX 12.8 and ran on a computer with Intel Xenon(R) CPU E5-2690 2.6GHz (2 Processors) and 20 GB of RAM. Metaheuristic algorithms and lower bounding techniques were coded in Matlab R2017 and ran on a computer with Intel(R) Core i5-4300U CPU @ 1.90 GHz and 8 GB of RAM.

6-1-Data generation

In this section, we explain data generation mechanism to evaluate the developed master schedules. To this end, we generated networks composed of $|H|$ echo departments, at the same quality level and reliable to each other in cooperation, while $|H| \in \{1,2,3,4,5\}$. This size of network is appropriate and applicable for setting up an echocardiography network between similar hospitals such as Specialty Cardiac Hospitals in Tehran. To generate realistic network instances, we adjusted the parameters based on the characteristics of THC. Each echo department of the generated networks is equipped with $|L_b|$ echo labs with different facilities while $|L_b|$ is randomly generated from set {7,8,9,10,11}. There are four types of echo labs: ω_1

(regular echo labs for performing echo types 1, 2 and 3), ω_2 (labs provided with specific facilities of echo type 4), ω_3 (labs provided with specific facilities of echo types 5) and ω_4 (labs provided with all facilities). Each lab is categorized in one of these types with the probability of about 71%, 15%, 7%, and 7%, respectively. A group of $|J|$ echocardiologists, fellows, residents and heart specialists with different specialty and quickness levels work at these networks such that $|J| = |L| + 3 \times |H|$. These specialists work at either one, two or even three centers. Specialists are categorized into three levels based on their specialty levels: ν_1 (specialists capable of performing echo type 1), ν_2 (specialists capable of performing echo types 1 and 2) and ν_3 (specialists capable of performing all echo types). Each specialist is in one of these types with the probability of about 35%, 35%, and 30%, respectively. The preferences of specialists for working in different shifts and hospitals are generated as a matrix with the density of 0.8, and nonzero values are generated as $unif[0.1,1]$. $unif[\,]$ stands for uniform distribution. Maximum and minimum allowable hours to work for each specialist are generated as $unif[26,30]$ and $unif[16,20]$ in hour, respectively. Total maximum allowable overtime for each specialist is generated as $unif[0,3]$ in hour.

We consider one week as the planning horizon. Each week consists of five days with two shifts (morning and evening) and one day with one shift (morning). The regular available times of echo labs in the morning and evening shifts are generated as $unif[3.5,5]$ and $unif[2,3.5]$ in hour, respectively. The maximum allowable overtime of echo labs in each shift is generated as $unif[0,1]$ in hour. Overtime cost per hour is assumed to be equal to 40.

Five major types of echos are performed in these networks. Demands of inpatients and outpatients for each echo type are generated in the interval reported in the first two columns of table 3. $|L|$ is the total number of labs. Also α is considered to be equal to 1, 1.05 and 1,1 to generate three different levels of demands. Generated values should be rounded to be integer. The percentage of demand that should necessarily be covered within this planning horizon, the revenue, and the expected duration of each echo type performed by different specialists are also generated in the intervals reported in table 3. It is worth noting that the values generated for duration of echo include the real time of echo and the required time for all measures just before and after echo such as preparation and writing of the report. For the sake of simplicity, we have normalized the values of revenue and cost. Regarding five levels for the number of hospitals, three levels for demand, and four test problems for each combination of them, totally, $3 \times 5 \times 4 = 60$ network instances were generated.

Table 3. Parameters of network instances

Echo type	Demand		Minimum percentage of covered demand		Revenue	Duration for different specialists (minute)
	Inpatients	Outpatients	Inpatients	Outpatients		
1	$\alpha\|L_b\|unif[11,16]$	$\alpha\|L\|unif[49,54]$	$unif[0.9,0.95]$	$unif[0.4,0.45]$	6	$unif[10,20]$
2	$\alpha\|L_b\|unif[24,29]$	$\alpha\|L\|unif[36,41]$	$unif[0.9,095]$	$unif[0.3,0.35]$	10	$unif[25,40]$
3	$\alpha\|L_b\|\ unif[0.15,0.3]$	$\alpha\|L\|unif[0.45,0.6]$	$unif[0.85,0.9]$	$unif[0.3,0.35]$	12	$unif[35,50]$
4	$\alpha\|L_b\|unif[0.15,0.3]$	$\alpha\|L\|unif[0.45,0.6]$	$unif[0.85,0.9]$	$unif[0.3,0.35]$	18	$unif[40,60]$
5	$\alpha\|L_b\|unif[1,1.6]$	$\alpha\|L\|unif[1.6,2.2]$	$unif[0.85,0.9]$	$unif[0.3,0.35]$	22	$unif[50,65]$

6-2-Parameter tuning

The fractional factorial experiment proposed by Taguchi (1986) was implemented to tune parameters. With the help of the signal-to-noise ratio (S/N) as a measure of variation, Taguchi method determines the best level of each parameter by conducting the minimum number of experiments. Sequence of neighborhood structures (SEQ), initial temperature of SA (T_0), cooling rate of SA (CR) and the percentage of time allocated to improvement phase (β) are parameters of VNS-SA that should be tuned. Likewise, number of particles (N_p), inertial weight (w), and the cognition learning factor (c_1) are

parameters of HPSO that should be tuned. When the value of c_1 is determined, the value of c_2 is calculated as $c_2 = 4.1 - c_1$. We used Minitab 18 to tune the parameters. Taguchi suggested L9 as the fittest orthogonal array design for both algorithms in two situations $|H| \in \{1,2,3\}$ and $|H| \in \{4,5\}$. Levels defined for each of these parameters and the best values for them are shown in table 4. Also S/N ratio plots are provided in Appendix D.

Table 4. Levels and the best values of parameters set by parameter tuning

	$\|H\| \in \{1,2,3\}$			$\|H\| \in \{4,5\}$		
	Parameter	Levels	Best value	Parameter	Levels	Best value
VNS-SA	*SEQ*	A,B,C	A	*SEQ*	A,B,C	A
	T_0	100,150,200	100	T_0	100,150,200	100
	CR	0.8,0.9,0.95	0.9	*CR*	0.8,0.9,0.95	0.9
	β	0.05,0.1,0.2	0.05	β	0.05,0.1,0.2	0.1
HPSO	N_p	10,20,30	10	N_p	5,10,15	5
	w	0.5,1,1.5	1	w	0.5,1,1.5	1
	c_1	1.5,2,2.5	2	c_1	1.5,2,2.5	2

Levels considered for the sequence of neighborhood structures in VNS-SA are defined as follows:

A: Neighborhood structures with minor changes precede those with major changes: $A = \{1-2-3-4-5-6-7-8-9-10\}$.

B: Neighborhood structures with major changes precede those with minor changes: $B = \{10-9-8-7-6-5-4-3-2-1\}$.

C: Neighborhood structures are in decreasing order of their effectiveness (based on their performance in solving several problems): $C = \{1-2-4-5-3-8-10-6-9-7\}$.

The values of other parameters of the algorithms, which are set according to the experience of solving many problems, are reported in table 5. Recall that $weight_1$, $weight_2$ and $weight_3$ are coefficients applied to find the initial solution of VNS-SA (section 4.1.2). The last three columns of this table are related to the models proposed for the initial solution of VNS-SA (section 4.1.2) and the fitness function of both algorithms (section 4.1.3). These models are not capable of reaching the optimality gap equal to zero in a short time, but they can find good and near optimal solutions so fast. Consequently, maximum computational times are required to be considered for them.

Table 5. Values of other parameters of algorithms, set by experience

	VNS-SA					HPSO
Parameter	($weight_1$, $weight_2$, $weight_3$)	Final temperature of SA	Number of iteration in each temperature	Time for initial solution	Time for fitness function	Time for fitness function
Values	(1000,700,20)	1	1	5 sec	1 sec	1 sec

6-3-Comparing the performance of metaheuristic algorithms

Using the mechanism explained in 6.1, we generated 60 test problems and then ran the MILP model, metaheuristic algorithms and lower bounding techniques for them. To obtain a lower bound and upper bound/optimal solution by the MILP model, maximum computational times were considered. These maximum computational times, depending on the number of hospitals, were from 2 hours (for $|H| = 1$) to

10 hours (for $|H| = 5$). Also for metaheuristic algorithms, maximum computational times, depending on the number of hospitals, from 1000 seconds (for $|H| = 1$) to 3600 seconds (for $|H| = 5$) were considered. The results are summarized in Appendix E. Values in the column of UB_{MILP} are the obtained upper bounds by the MILP after the defined maximum computational time, or when the procedure stops with the "*out of memory*" error before this time. Also "-" in this column implies that no feasible solution could be found in defined maximum computational time or by the time of *out of memory* error. All the comparisons and analysis performed in this section and the next sections are based on the results provided in Appendix E.

About the computational time of metaheuristics, two notes worth to be mentioned: First, since the construction of CMES for a network of several hospitals requires considering a large number of echo labs, specialists and patients, and also due to the structure of the developed algorithms, which contain a mathematical model, we defined this maximum computational time empirically such that a high quality solution is obtainable. Second, the developed CMES does the shift scheduling and capacity allocation in a network of hospitals for a planning horizon of one week, which can be repeated for several weeks (usually for a month); thus, the defined computational times are totally acceptable.

In order to compare the performance of metaheuristic algorithms, the results of VNS-SA and HPSO were normalized by equation. (62).

$$RPD_{i,j} = \frac{\left(sol_{i,j} - sol_{i,best}\right)}{sol_{i,best}} \times 100 \tag{62}$$

Where, i is the index of each test problem, and j is the index of each algorithm. $sol_{i,j}$ and $sol_{i,best}$ are the solution of test problem i with algorithm j, and the best obtained solution of test problem i, respectively. Then a paired-t test at the 95% significance level was applied. The hypothesis test is stated as:

$$\begin{cases} H_0: \ \mu_{VNS-SA} = \mu_{HPSO} \\ H_1: \ \mu_{VNS-SA} \neq \mu_{HPSO} \end{cases}$$

The results are provided in table 6. According to this table, P-value is equal to 0.000 showing that there is a significant difference between the performances of the two algorithms. The calculated average objective function obtained by these algorithms indicates that VNS-SA outperforms HPSO. Although HPSO had the advantage of using the priority table instead of the specialists' assignment table to decrease complexity, it could not overcome VNS-SA in quality of solutions. We believe that this is due to the weak performance of the algorithm proposed for creating specialists' assignment table out of the priority table (Section 4.2.3), and there is space to improve it as a future research.

Table 6. Statistical comparison of VNS-SA and HPSO

Average objective function by VNS-SA	Average objective function by HPSO	Confidence interval (95%)	T-value	P-value
32441.78	31875.25	(-1.876, -1.354)	-12.39	0.000

6-4-Comparing the performance of the best metaheuristic algorithm (VNS-SA) with the upper bound obtained by the mathematical model

In this section, we analyze the results to prove the verification and efficiency of VNS-SA. Table 7 summarizes the results of MILP and VNS-SA in each level of number of hospitals on the test problems for which MILP has stopped with a feasible solution. While VNS-SA has been able to find acceptable

solutions for all the test problems, MILP has failed to find feasible solutions for two, six, seven, eight and nine of the test problems with one, two, three, four and five hospitals, respectively. In other words, for 53% of the test problems, MILP has stopped after the predefined computational time or due to the *out of memory* error without any feasible solution. For the test problems reflecting our idea of centralized scheduling in several hospitals (i.e., test problems with $|H| > 1$), MILP has failed to find feasible solutions for 62% of the test problems. As the number of variables and constraints increase by considering more hospitals, the capability of MILP for finding feasible solution decreases. These results prove the dominant performance of VNS-SA over MILP, and the necessity of applying VNS-SA instead of MILP for centralized appointment scheduling systems.

MILP and VNS-SA could also be compared in terms of the time they require to find the first feasible solution (for those problems that MILP has stopped with a feasible solution). While VNS-SA found a feasible solution in a quite short time (the considered time for creating initial solution), it took a long time for MILP to find a feasible solution in most problems with number of hospitals greater than two. For example, the first feasible solution for test problem 35 was obtained after about 3000 seconds by MILP. However, we are not going to compare MILP and VNS-SA from this point of view here. As a matter of fact and according to the experiments, MILP requires a large amount of time to find a feasible solution for this problem and to improve it. The reason why we considered large computational time for MILP lies in this fact and with the hope of providing feasible solutions for more problems. Although even considering this large computational time, no feasible solutions were found for the majority of test problems.

Regarding 28 test problems for which MILP has stopped with a feasible solution, the last column of Appendix E shows the gap between the upper bounds of MILP and the solutions of VNS-SA. According to this column, although in 12 cases out of these 28 test problems MILP has provided better solutions, the average gap for these cases is only 0.57%, and the maximum gap is less than 1.3%. As shown in table 7, the average solutions of MILP, VNS-SA and the average gap between them for these 28 test problems show that their solutions do not differ considerably. Based on the solutions and the considered computational times, it can be concluded that VNS-SA can obtain a solution as good as the upper bound of MILP in considerably less computational time. These results verify that VNS-SA is valid, reliable, and from computational time aspect, reasonable.

Table 7. Results of MILP and VNS-SA on problems for which MILP has stopped with a feasible solution

$\|H\|$	#FS*	Average upper bound of MILP (UB_{MILP})	Average solution of VNS-SA (Sol_{VNS-SA})	Average Gap% (Sol_{VNS-SA} vs. UB_{MILP})**
1	10	10215.6	10255.6	0.38
2	6	22237.17	22350.67	0.52
3	5	34685.4	34720.8	0.12
4	4	43782.75	44414.75	1.60
5	3	55530	55041.33	-0.88
average		26811.68	26894.54	0.40

*Number of problems for which MILP has stopped with a feasible solution

** Gap (Sol_{VNS-SA} vs. UB_{MILP}) = $\frac{(Sol_{VNS-SA} - UB_{MILP})}{UB_{MILP}} \times 100$

6-5-Comparing the performance of lower bounding techniques

In order to compare the performance of lower bounding techniques, the results of LB_1 and LB_2 (provided in Appendix E) were normalized by equation (62). Considering the following hypothesis test, a paired-t test at the 95% significance level was applied.

$$\begin{cases} H_0: \ \mu_{LB1} = \mu_{LB2} \\ H_1: \ \mu_{LB1} \neq \mu_{LB2} \end{cases}$$

Table 8. Statistical comparison of LB_1 and LB_2

Average lower bound by LB_1	Average lower bound by LB_2	Confidence interval (95%)	T-value	P-value
36365.67	47437.03	(-35.44 , -29.69)	-22.76	0.000

According to table 8, since P-value is equal to 0.000, there is a significant difference in the performance of lower bounding techniques. Better average lower bound provided by LB_1, as reported in Table 8, shows the superiority of LB_1 over LB_2. The reason seems to be the fact that LB_2 does not consider demands. Thus, it increases the number of performed echoes with the lowest durations and the highest revenues without limitation. Also from computational time point of view, while LB_1 presents the lower bound in less than 0.1 second, LB_2 cannot reach the optimality gap equal to zero in a short time, so a limited computational time like 10 seconds should be considered for it.

6-6-Comparing the performance of the best lower bounding technique (LB_1) with the lower bound obtained by the mathematical model

As it can be seen in Appendix E, LB_1 technique has provided better lower bounds for 34 instances out of 60 generated test problems compared to MILP. For the rest of the test problems, both LB_1 and MILP have provided equal lower bounds. To investigate if this difference is statistically significant, a paired-t test at the 95% significance level was applied. The hypothesis test is as follows:

$$\begin{cases} H_0: \ \mu_{LB_{MILP}} = \mu_{LB1} \\ H_1: \ \mu_{LB_{MILP}} \neq \mu_{LB1} \end{cases}$$

Table 9. Statistical comparison of LB_1 and $MILP$

Average lower bound by $MILP$	Average lower bound by LB_1	Confidence interval (95%)	T-value	P-value
36804.58	36365.67	(-4.678 , -2.586)	-6.95	0.000

With respect to P-value, as well as the average lower bound of MILP and LB_1, reported in table 9, it is concluded that LB_1 has significantly better performance in providing lower bounds for the problem. Also from computational time point of view, while the lower bounds provided by MILP are obtained after the considered large computational time, LB_1 presents its lower bound in less than 0.1 second for all the test problems.

6-7-Managerial insights and model extension ideas

Urgent cases: Since in many hospitals such as THC, an echo lab exists in the Emergency Department, we did not take urgent cases into account in our approaches. However, the structures of the proposed approaches let us incorporate them easily into the model by redefining set I such that each index i indicates both echo type and urgency condition. In addition, regarding W_{ib}, while these coefficients can be simply defined as revenue of each echo type, they can be defined in many other ways such as weight of each echo type. Thus we can easily consider urgent patients with higher values of W_{ib} in our approaches.

***Higher priorities of inpatients*:** In the current situation of the observed echo departments, sometimes, high demand of outpatients may cause the schedulers ignore demands of inpatients; this makes the length of stay much longer, causes bed shortage and disrupts patient flow. Appropriate values for W_{ib} can give

higher priorities to inpatients to manage this situation. Also, by considering higher values for K_{ib} for inpatients, more percentage of demand of inpatients can be covered in the planning horizon.

***Educational purposes*:** Sometimes, for educational purposes, some specialists are required to perform a certain number of echos of each type. In order to incorporate this assumption into the model and our best metaheuristic approach (VNS-SA), the following constraint should be added to the MILP model and the fitness function of metaheuristic. Q_{ij} is a parameter that specifies the minimum number of echos of type i that specialist j should perform in the planning horizon.

$$\sum_{t\in T, b\in B, l\in L} X_{iblt}\, \mathrm{Y}_{jlt} \geq Q_{ij} \qquad \forall i \in I, j \in J \qquad (63)$$

Since this constraint is nonlinear, it can be substituted by the following linear set of constraints in both MILP model and the fitness function of metaheuristic. X'_{ibjt} is a variable that determines the number of echos of type i in situation b performed by specialist j in shift t. Constraint set (64) ensures that if a specialist is not assigned to any lab in a specific shift, the total number of echos performed by that specialist in that shift is equal to zero. Constraint set (65) guarantees that the number and type of echoes performed by any specialist in any shift are equal to the number and type of performed echoes in the lab to which that specialist in that shift is assigned. Constraint set (66) assures that each specialist performs the predefined minimum number of echos of each type. Finally, Constraint set (67) defines the required integer variable.

$$\sum_{b\in B, t\in T} X'_{ibjt} \leq M \sum_{l\in L} \mathrm{Y}_{jlt} \qquad t \in T, j \in J \qquad (64)$$

$$X'_{ibjt} \leq X_{iblt} + M\left(1 - Y_{jlt}\right) \qquad \forall i \in I, b \in B, l \in L, t \in T, j \in J \qquad (65)$$

$$\sum_{b\in B, t\in T} X'_{ibjt} \geq Q_{ij} \qquad \forall i \in I, j \in J \qquad (66)$$

$$X'_{ibjt} \geq 0, int \qquad \forall i \in I, t \in T, j \in J, b \in B \qquad (67)$$

***Specialists' preferences*:** To consider specialists' preferences more, a second objective function to maximize the total satisfied specialists' preferences or the following constraint can be added to the MILP model. Constraint set (68) obligates the model to satisfy a certain level of total satisfaction, referred to as FT.

$$\sum_{j\in J, t\in T, l\in L, h\in H} F_{jth} N_{lh} \mathrm{Y}_{jlt} \geq FT \qquad (68)$$

Moreover, in order to incorporate this assumption into VNS-SA, two adjustments are required: Constraint set (68) should be added to the initial solution construction model in section 4.1.2, and condition 3 should be added to the conditions mentioned in section 4.1.4.

- Condition 3: The sum of satisfied preferences of specialists should be greater than FT.

To show the applicability of two recent model extension ideas, the results of five test problems are reported in table 10. Note that there is no need to change the LB_1 technique. These results verify the great performance of proposed metaheuristic and lower bounding technique for the discussed problem with mentioned extension ideas.

Table 10. Results of model extension ideas

H	$UB_{extended\_MILP}$	$LB_{extended\_MILP}$	$Sol_{extended\_VNS-SA}$	LB_1
1	8901	10806	9083	10091
2	_	26975	22661	26098.4
3	_	36006	30162	33452.8
4	_	44270	40642	44270
5	_	65108	53600	65057

7-Conclusions

In the present study, we approached the appointment scheduling problem in a distributed network of echocardiography departments with the objective function of maximization of profit margin through maximizing the number of performed echoes and minimizing overtime. In this network, various types of inpatients/outpatients request for different types of echos. Each echo is performed by a specialist with related specialty level and in a lab with required facilities. In order to develop an efficient centralized master schedule for echocardiography networks, which we called CMES, we proposed approaches that are able to handle shift scheduling and capacity allocation problems simultaneously, considering the possibility of overtime. First, we developed a MILP model. This model requires a large amount of time and memory to solve the problem; however, in most cases, it cannot find any feasible solution, specialy when the size of the network increases. This being the case, two metaheuristic algorithms (VNS-SA and HPSO) with different approaches were presented to find good quality solutions in acceptable computational time. The novelty of the developed metaheuristics is that they only search within the feasible solution area of shift scheduling, and a mathematical model, as the fitness function of metaheuristics, solves the obtained capacity allocation problem. Furthermore, two lower bounding techniques (LB_1 and LB_2) were developed. Based on the conducted experiments, VNS-SA was shown to present a better performance comparing to HPSO in terms of quality of solutions. Then, the efficiency and validity of VNS-SA were examined and confirmed by comparing the performance of VNS-SA and MILP. According to the obtained results, MILP is unable to find the solution for 62% of networks with more than one echo department, while VNS-SA finds good quality solutions for all the problems in reasonable amount of time. In addition, the performance of VNS-SA was verified due to the acceptable deviation from the solutions found by MILP (in those problems for which MILP stopped with a feasible solution). Next, the lower bounding techniques were evaluated. It was concluded that LB_1 is able to provide better lower bounds in comparison with LB_2 and the mathematical model. The reason of the superiority of LB_1 over LB_2 is that it incorporates more effective factors. Finally, managerial insights and some ideas for the algorithm and model extension were discussed.

Lastly, several applicable areas for further research can be suggested. First, decisions in the operational level of decision making such as determining the sequence and exact appointment time for patients, aiming at minimization of patients' waiting time and specialists' idle-time can be considered. Second, HPSO can be enhanced in order to take advantage of its main ideas, which were eliminating the need for complicating neighborhood structures and getting rid of infeasible specialists' assignments. To do so, we suppose that improving the algorithm proposed for creating specialists' assignment tables out of priority tables is beneficial. Third, stochastic or robust approaches can be developed in case of high variability (in demand or duration of echo) or any similar settings in which rescheduling is critical.

References

about Tehran Heart Center. (2017). Retrieved from http://thc.tums.ac.ir/default.aspx

Ahmadi-Javid, A., Jalali, Z., & Klassen, K. J. (2017). Outpatient appointment systems in healthcare: A review of optimization studies. *European Journal of Operational Research, 258*(1), 3-34.

Aringhieri, R., Landa, P., Soriano, P., Tànfani, E., & Testi, A. (2015). A two-level metaheuristic for the operating room scheduling and assignment problem. *Computers & Operations Research, 54*, 21-34.

Bailey, N. T. (1952). A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times. *Journal of the Royal Statistical Society. Series B (Methodological)*, 185-199.

Bakker, M., & Tsui, K.-L. (2017). Dynamic resource allocation for efficient patient scheduling: a data-driven approach. *Journal of Systems Science and Systems Engineering, 26*(4), 448-462.

Bakshi, S. (2013). Business process re-engineering a cardiology department. *World hospitals and health services: the official journal of the International Hospital Federation, 50*(2), 40-45.

Bettinelli, A., Cacchiani, V., & Malaguti, E. (2017). A branch-and-bound algorithm for the knapsack problem with conflict graph. *INFORMS Journal on Computing, 29*(3), 457-473.

Brech, C.-H., Ernst, A., & Kolisch, R. (2019). Scheduling medical residents' training at university hospitals. *European Journal of Operational Research, 274*(1), 253-266.

Cabell, C. H., Trichon, B. H., Velazquez, E. J., Dumesnil, J. G., Anstrom, K. J., Ryan, T., Miller, A. B., Belkin, R. N., Cropp, A. B., & O'connor, C. M. (2004). Importance of echocardiography in patients with severe nonischemic heart failure: the second Prospective Randomized Amlodipine Survival Evaluation (PRAISE-2) echocardiographic study. *American heart journal, 147*(1), 151-157.

Castro, E., & Petrovic, S. (2012). Combined mathematical programming and heuristics for a radiotherapy pre-treatment scheduling problem. *Journal of Scheduling, 15*(3), 333-346.

Cayirli, T., & Veral, E. (2003). Outpatient scheduling in health care: a review of literature. *Production and operations management, 12*(4), 519-549.

Gandhi, R. (2013). *Increasing the Daily Throughput of Echocardiogram Patients using Discrete Event Simulation.* (Master of Health Science), University of Toronto.

Geronimo, R. (2017). *Improving Stress Echocardiogram Access for Patients with Low-Risk Chest Pain in the Emergency Department Clinical Decision Unit.* (Master's thesis), University of San Francisco.

Hamid, M., Barzinpour, F., Hamid, M., & Mirzamohammadi, S. (2018). A multi-objective mathematical model for nurse scheduling problem with hybrid DEA and augmented ε-constraint method: A case study. *Journal of Industrial and Systems Engineering, 11(Special issue: 14th International Industrial Engineering Conference)*, 98-108.

Hamid, M., Nasiri, M. M., Werner, F., Sheikhahmadi, F., & Zhalechian, M. (2019). Operating room scheduling by considering the decision-making styles of surgical team members: a comprehensive approach. *Computers & Operations Research, 108*, 166-181.

Heath, S. (2017). Centralized Appointment Scheduling Aids Patient Experience. Retrieved from https://patientengagementhit.com/news/centralized-appointment-scheduling-aids-patient-experience

Hong, Y.-C., Cohn, A., Epelman, M. A., & Alpert, A. (2018). Creating resident shift schedules under multiple objectives by generating and evaluating the Pareto frontier. *Operations Research for Health Care,*. doi:https://doi.org/10.1016/j.orhc.2018.08.001.

Hulshof, P. J., Kortbeek, N., Boucherie, R. J., Hans, E. W., & Bakker, P. J. (2012). Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS. *Health systems, 1*(2), 129-175.

Katsi, V. K., Vrachatis, D. A., Politi, A., Papageorgiou, M., Koumoulidis, A., Vlasseros, I., Vavuranakis, M., Tousoulis, D., Stefanadis, C., & Kallikazaros, I. (2014). Cardiac echo-lab productivity in times of economic austerity. *SpringerPlus, 3*(1), 703.

Kennedy, j., & Eberhart, R. (1995). *Particle swarm optimization.* Paper presented at the Proceeding IEEE Inter Conference on Neural Networks, Perth, Australia, Piscat-away.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simmulated annealing. *science, 220*(4598), 671-680.

Klassen, K. J., & Yoogalingam, R. (2019). Appointment scheduling in multi-stage outpatient clinics. *Health Care Management Science, 22*(2), 229-244.

LaGanga, L. R., & Lawrence, S. R. (2012). Appointment overbooking in health care clinics to improve patient service and clinic performance. *Production and operations management, 21*(5), 874-888.

Munt, B., O'Neill, B., Koilpillai, C., Gin, K., Jue, J., & Honos, G. (2006). Treating the right patient at the right time: Access to echocardiography in Canada. *Canadian Journal of Cardiology, 22*(12), 1029-1033.

Murray, M., & Berwick, D. M. (2003). Advanced access: reducing waiting and delays in primary care. *Jama, 289*(8), 1035-1040.

Nazif, H. (2018). Operating Room Surgery Scheduling with Fuzzy Surgery Durations Using a Metaheuristic Approach. *Advances in Operations Research, 2018*. doi:https://doi.org/10.1155/2018/8637598

Nguyen, T. B. T., Sivakumar, A. I., & Graves, S. C. (2015). A network flow approach for tactical resource planning in outpatient clinics. *Health Care Management Science, 18*(2), 124-136.

Parallon. (2013). Using Centralized Scheduling to Improve Performance. Retrieved from http://go.parallon.com/hubfs/WebContent/ParallonRevenueCycle_CaseStudy_CentralizedScheduling.pdf

Pena, S. M., & Lawrence, N. (2017). Analysis of wait times and impact of real-time surveys on patient satisfaction. *Dermatologic Surgery, 43*(10), 1288-1291.

Penn, M., Potts, C. N., & Harper, P. R. (2017). Multiple criteria mixed-integer programming for incorporating multiple factors into the development of master operating theatre timetables. *European Journal of Operational Research, 262*(1), 194-206.

Poli, R., Kennedy, J., & Blackwell, T. (2007). Particle swarm optimization. *Swarm intelligence, 1*(1), 33-57.

Rabani, M., & Niyazi, M. (2017). Solving a nurse rostering problem considering nurses preferences by graph theory approach. *Journal of Industrial and Systems Engineering, 10(special issue on scheduling)*, 38-57.

Reliasmedia. (2014). Centralized scheduling provides many benefits. Retrieved from https://www.reliasmedia.com/articles/16945-centralized-scheduling-provides-many-benefits

Roshanaei, V., Luong, C., Aleman, D. M., & Urbach, D. (2017). Propagating logic-based Benders' decomposition approaches for distributed operating room scheduling. *European Journal of Operational Research, 257*(2), 439-455.

Santibáñez, P., Begen, M., & Atkins, D. (2007). Surgical block scheduling in a system of hospitals: an application to resource and wait list management in a British Columbia health authority. *Health Care Management Science, 10*(3), 269-282.

Saure, A., Patrick, J., Tyldesley, S., & Puterman, M. L. (2012). Dynamic multi-appointment patient scheduling for radiation therapy. *European Journal of Operational Research, 223*(2), 573-584.

Scisolutions. (2017). centralized patient access: a checklist for success. Retrieved from https://www.scisolutions.com/uploads/data-sheets/Checklist_Centralized-PtAccess-061317F.pdf

Smalley, H. K., & Keskinocak, P. (2016). Automated medical resident rotation and shift scheduling to ensure quality resident education and patient care. *Health Care Management Science, 19*(1), 66-88.

Taguchi, G. (1986). *Introduction to quality engineering: designing quality into products and processes*.

Tohidi, M., Kazemi Zanjani, M., & Contreras, I. (2019). Integrated physician and clinic scheduling in ambulatory polyclinics. *Journal of the Operational Research Society, 70*(2), 177-191.

Trang, A., Kampangkaew, J., Fernandes, R., Tiwana, J., Misra, A., Hamzeh, I., Blaustein, A., Aguilar, D., Shah, T., & Ballantyne, C. (2019). Understanding by General Providers of the Echocardiogram Report. *The American journal of cardiology, 124*(2), 296-302.

Tsai, P.-F. J., & Teng, G.-Y. (2014). A stochastic appointment scheduling system on multiple resources with dynamic call-in sequence and patient no-shows for an outpatient clinic. *European Journal of Operational Research, 239*(2), 427-436.

Vali Siar, M., Gholami, S., & Ramezanian, R. (2017). Multi-period and multi-resource operating room scheduling and rescheduling using a rolling horizon approach: a case study. *Journal of Industrial and Systems Engineering, 10(special issue on healthcare)*, 97-115.

Vogl, P., Braune, R., & Doerner, K. F. (2019). Scheduling recurring radiotherapy appointments in an ion beam facility. *Journal of Scheduling, 22*(2), 137-154.

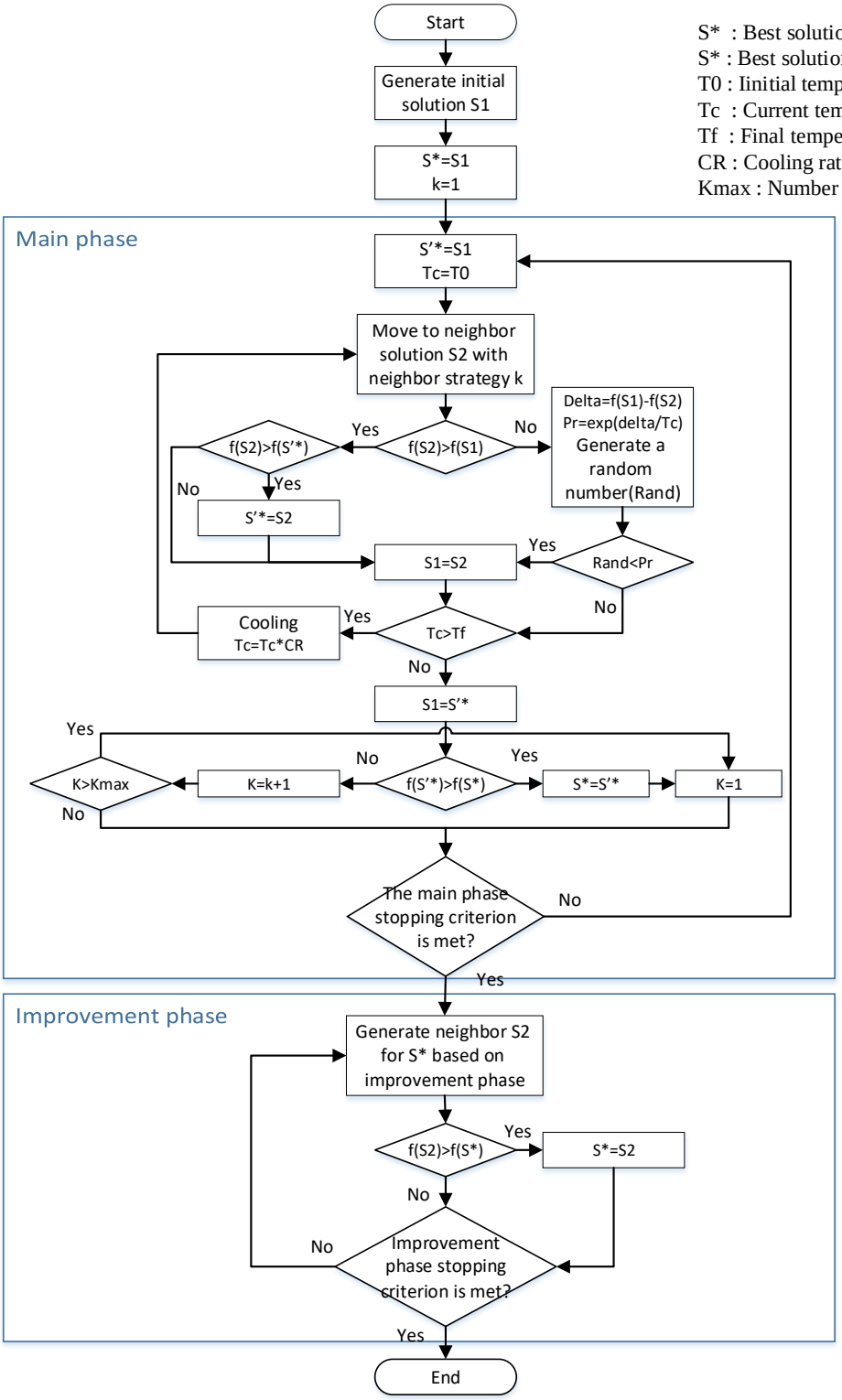
Volland, J., Fügener, A., & Brunner, J. O. (2017). A column generation approach for the integrated shift and task scheduling problem of logistics assistants in hospitals. *European Journal of Operational Research, 260*(1), 316-334.

Welch, J., & Bailey, N. J. (1952). Appointment systems in hospital outpatient departments. *The lancet, 259*(6718), 1105-1108.

Zhang, P., Bard, J. F., Morrice, D. J., & Koenig, K. M. (2019). Extended open shop scheduling with resource constraints: Appointment scheduling for integrated practice units. *IISE Transactions, 51*(10), 1-24.

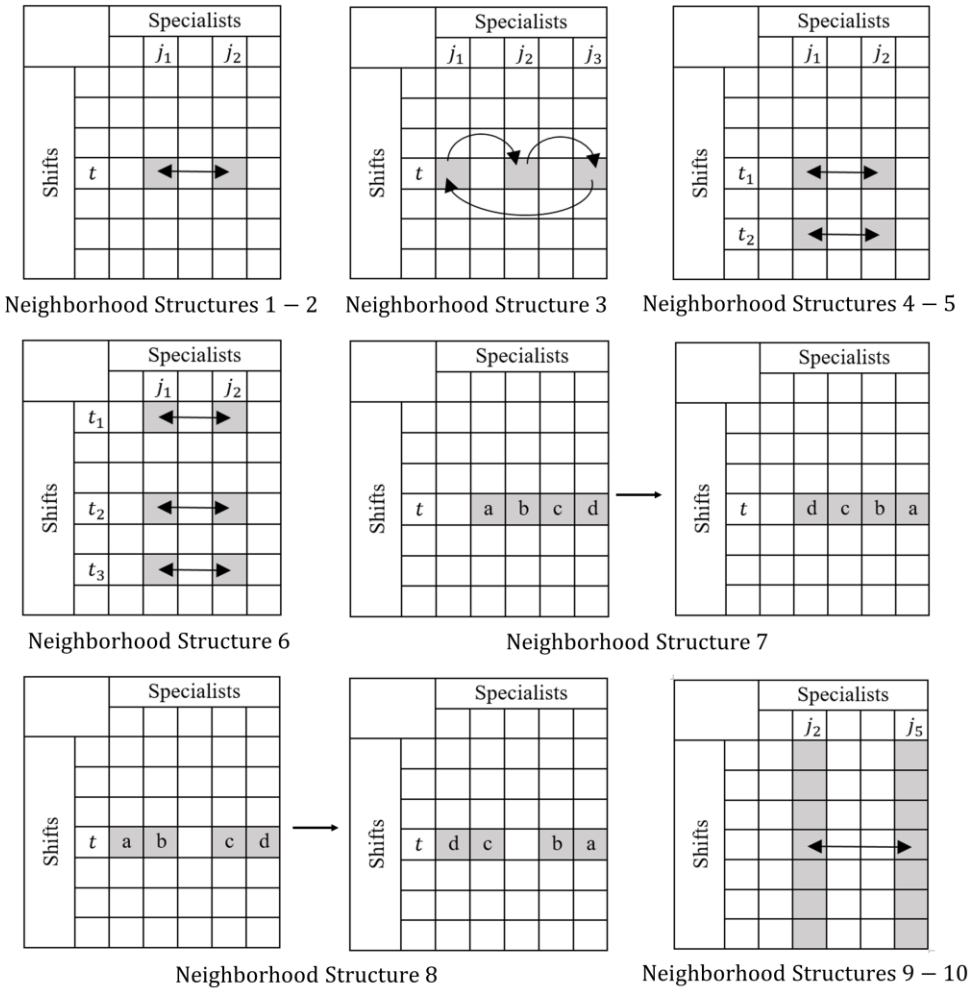
Zhou, L., Geng, N., Jiang, Z., & Wang, X. (2017). Combining revenue and equity in capacity allocation of imaging facilities. *European Journal of Operational Research, 256*(2), 619-628.
Zhou, L., Geng, N., Jiang, Z., & Wang, X. (2018). Multi-objective capacity allocation of hospital wards combining revenue and equity. *Omega, 81*, 220-233.

Appendix A: Flowchart of VNS-SA

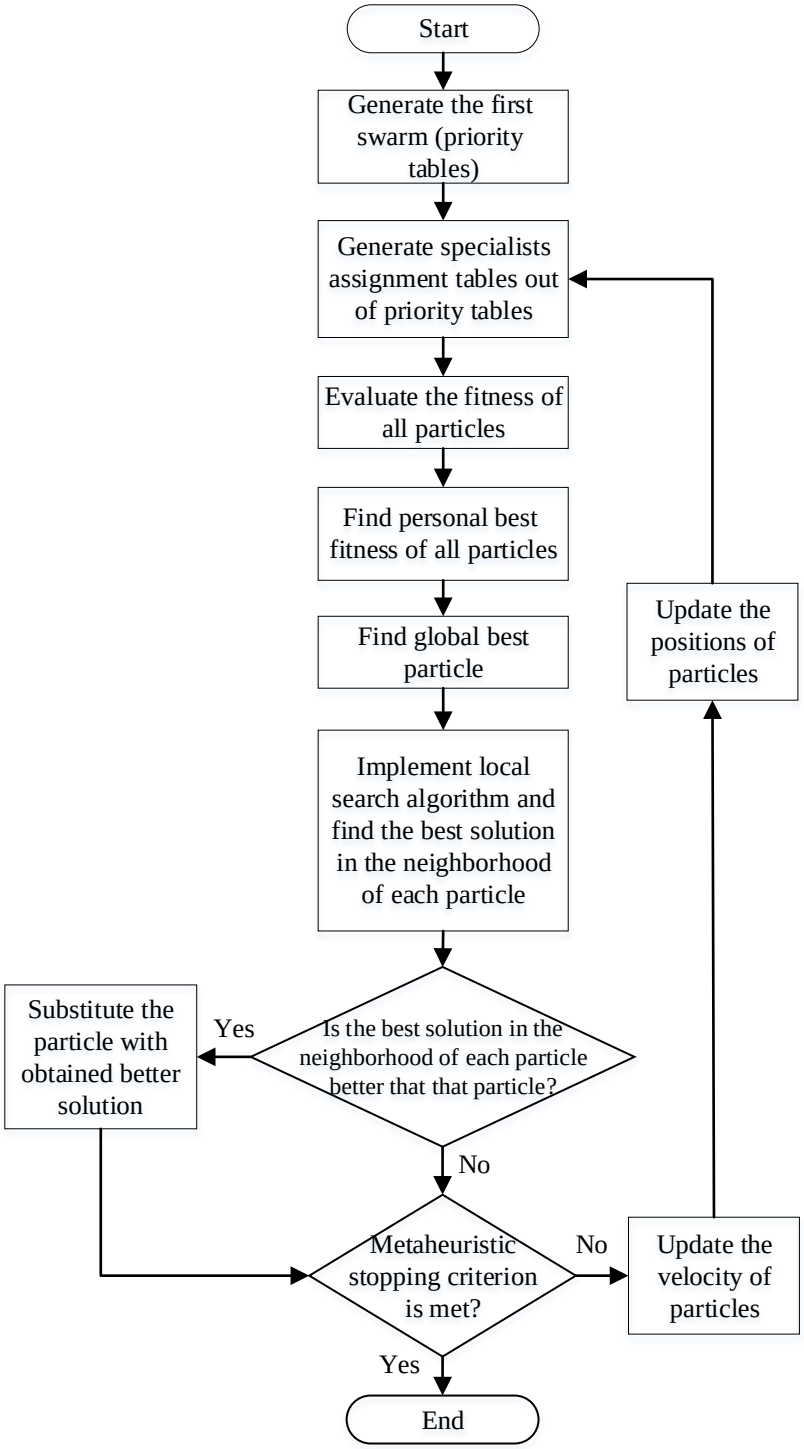


S* : Best solution
S* : Best solution of each iteration
T0 : Iinitial temperature
Tc : Current temperature
Tf : Final temperature
CR : Cooling rate
Kmax : Number of neighborhood structures

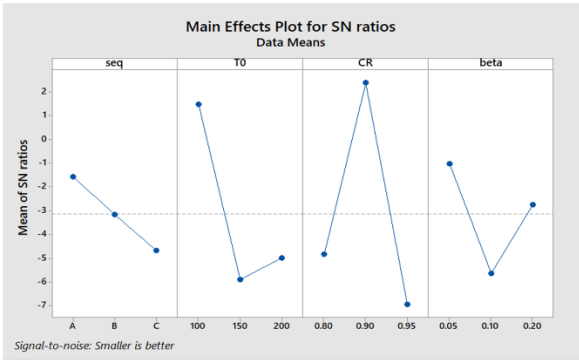
Appendix B: Neighborhood structures of VNS-SA



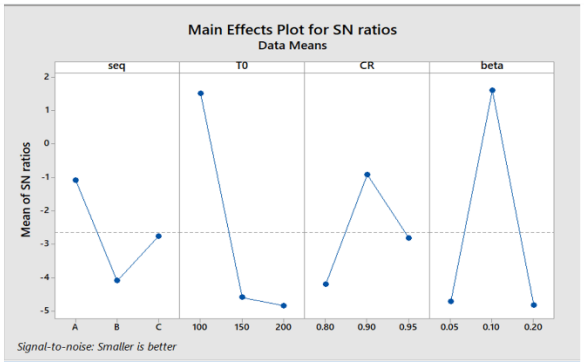
Appendix C: Flowchart of HPSO



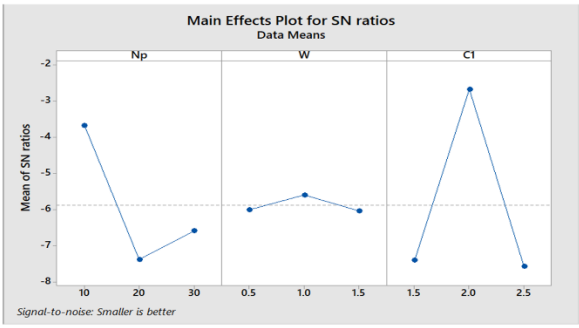
Appendix D : The S/N ratio plots



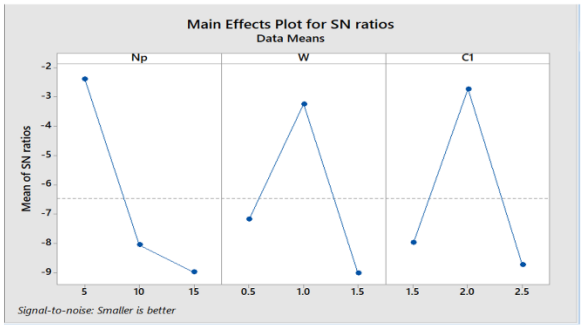
a. VNS − SA , $|H| \in \{1,2,3\}$



b. VNS − SA , $|H| \in \{4,5\}$



c. HPSO , $|H| \in \{1,2,3\}$



d. HPSO , $|H| \in \{4,5\}$

Appendix E : Results of model, metaheuristic algorithms and lower bounding techniques

$\|H\|$	Level of demand	#Test problem	Mathematical model			Metaheuristic algorithms		Lower bounding technique		Gap % (Sol_{VNS-SA} vs. UB_{MILP})*
			UB_{MILP}	LB_{MILP}	Gap % (UB_{MILP} vs. LB_{MILP})	Sol_{VNS-SA}	Sol_{HPSO}	LB_1	LB_2	
1	1	1	7962	9278	16.53	7946	7828	9056.6	11625.8	-0.20
		2	7980	9170	14.91	7998	7925	8718.66	11679.4	0.23
		3	10114	11578	14.47	10233	10001	11578	14268.82	1.18
		4	13174	13878	5.34	13218	13074	13878	17544.87	0.33
	2	5	10010	11690	16.78	10026	9984	10528	17556.56	0.16
		6	13723	15691	14.34	13613	13469	15005	18886.1	-0.80
		7	9520	11296	18.66	9544	9512	10501.6	16640.9	0.25
		8	10993	11805	7.39	11049	10949	11805	17121.23	0.51
	3	9	-	14427	-	11479	11387	13280.6	19980.36	-
		10	-	11821	-	9381	9341	11821	14695.35	-
		11	7446	8984	20.66	7440	7424	8028	13160.08	-0.08
		12	11234	13454	19.76	11489	11146	12541.2	18037.87	2.27
2	1	13	-	21777	-	18485	17980	20134	28756.66	-
		14	21660	23892	10.30	21688	21396	23892	30947.65	0.13
		15	20190	23419	15.99	20283	20067	21297.8	31857.53	0.46
		16	24055	26975	12.14	24083	23685	26975	32741.97	0.12
	2	17	-	22132	-	18892	18664	21136.8	28602.6	-
		18	-	26303	-	22483	22331	26183	32615.43	-
		19	-	24928	-	20632	20467	22228.8	29118.03	-
		20	19259	22445	16.54	19389	19197	21402.2	27077.85	0.68
	3	21	-	26582	-	21658	21454	25424	32522.96	-
		22	22729	27279	20.02	23055	22799	27279	31320.07	1.43
		23	25530	29002	13.60	25606	25226	27191.6	33873.93	0.30
		24	-	27189	-	22933	22297	24058.3	31823.51	-
3	1	25	-	33335	-	29715	29243	33335	43616.09	-
		26	40455	47836	18.24	40297	39532	43550.4	53901.6	-0.39
		27	-	32391	-	29083	28071	32008.6	44551.02	-
		28	34959	36887	5.52	34851	33964	36887	50490.7	-0.31
	2	29	-	39041	-	32829	32167	38436.4	50419.18	-
		30	-	32290	-	27826	27346	32290	45658.72	-
		31	29139	33728	15.75	29023	28833	31916	43170.88	-0.40
		32	36754	42700	16.18	36834	36098	40339.6	48935.1	0.22
	3	33	-	37419	-	31721	31295	37419	46995.2	-
		34	-	35483	-	31346	30659	35483	45129.91	-
		35	32120	38224	19.00	32599	32191	34138.6	43827.96	1.49
		36	-	42104	-	36444	35792	42104	49442.87	-
4	1	37	-	44788	-	37562	37443	42797.2	63349.83	-
		38	45044	51700	14.78	44933	44158	49541	68185.7	-0.25
		39	-	38721	-	38713	38342	38721	54633.33	-
		40	42635	50601	18.68	42117	41514	48468	64192.9	-1.21
	2	41	-	45115	-	42494	41336	45115	60772.21	-
		42	-	46338	-	43105	42507	46338	64594.76	-

$\|H\|$	Level of demand	#Test problem	Mathematical model			Metaheuristic algorithms		Lower bounding technique		Gap % (Sol_{VNS-SA} vs. UB_{MILP})*
			UB_{MILP}	LB_{MILP}	Gap % (UB_{MILP} vs. LB_{MILP})	Sol_{VNS-SA}	Sol_{HPSO}	LB_1	LB_2	
		43	40672	44294	20.78	44126	43154	44294	63805.78	8.49
		44	46780	57360	22.62	46483	46011	54831	68009.2	-0.63
	3	45	-	47160	-	42349	41333	47160	59708.45	-
		46	-	42873	-	36646	36491	42873	55109.34	-
		47	-	59361	-	48177	47981	56793	69446.2	-
		48	-	57909		43948	42965	55352	63612.2	-
5	1	49	-	51736	-	46083	44238	51736	69435	-
		50	55002	66842	21.53	54400	53501	63870	77896.7	-1.09
		51	-	64258	-	52898	51920	57913.6	81593	-
		52	55029	66672	21.16	54624	52570	63753	78291	-0.74
	2	53	-	56895	-	49407	49417	56895	78283	-
		54	-	57108	-	52650	52332	57108	75390	-
		55	-	58353	-	55103	54364	58353	78582	-
		56	56559	70562	24.76	56100	55090	67502	84076.9	-0.81
	3	57	-	66014	-	59040	57291	66014	88242	-
		58	-	63685	-	53934	52934	63685	77317.6	-
		59	-	66692	-	60952	60268	66692	86764.08	-
		60	-	79322	-	65492	62561	74283.6	86335.8	-

* Gap (Sol_{VNS-SA} vs. UB_{MILP}) = $\frac{(Sol_{VNS-SA}-UB_{MILP})}{UB_{MILP}} \times 100$