Scheduling of periodic services to customers in dispersed locations from heterogeneous multi-agent companies considering uncertainty: A real case study

H. Jafar-Zanjani a, M. Zandieh b,*, M. Khalilzadeh a

a Industrial Engineering Department, Science and Research Branch, Islamic Azad University, Tehran, Iran

b Department of Industrial Management, Management and Accounting Faculty, Shahid Beheshti University, G.C., Tehran, Iran

Abstract

The scheduling problem of periodic services from service providers to customers located in different places and need different services. The service centers are also located in different positions, each of which has limited number of teams with the capability of performing one or some services. The goal is to simultaneously minimize ‘total service costs’ and ‘total earliness/tardiness’ in providing services to customers. Providing an optimal maintenance schedule is a big challenge in those companies with dispersed supply centers. In this paper, a novel bi-objective mixed integer linear programming model along with augmented epsilon constraint method is presented to exactly solve this problem. Then, a bi-objective meta-heuristic technique based on genetic algorithm is proposed and its performance in solving large-scale problems is assessed. The uncertain parameters are faced through robust possibilistic programming approach to diminish the risk of decision making. Finally, the performance of the proposed model and solution approaches are evaluated through a real case study in maintenance scheduling of compressed natural gas (CNG) stations equipment in Iran.

Keywords: Scheduling; Bi-objective Optimization; Robust Possibilistic Programming; Genetic Algorithm; Uncertainty; Augmented Epsilon Constraint

1. Introduction

To be active and compete in the global business market, the companies are transforming from a centralized structure to a decentralized one. That is, considering the geographical dispersion of their customers, these companies establish representatives or similar companies in different places to provide them with more suitable services with a lower cost and waiting time. Such systems are called multi-factory production (MFP) [1].

These product supply centers or public service centers, which consist of several machines and equipment, are dispersed in several places to be available in less time and with lower cost [2]
One of the main issues about dispersed service centers is the repair and maintenance planning of facilities and equipment to increase their reliability and availability and also provide the customers with the required services or products in each time interval with a lower tardiness.

However, the failure of the components is an inevitable fact related to the production and service systems. These failures may be the result of inadequate testing and inspection, poor maintenance, human error, etc. Applying more efficient equipment can result in an improved system productivity and profitability, which relies intensively on the reliable maintenance strategies in a system [3, 4].

Generally, to improve the system reliability, implementing the optimal service plans is important to managers of a company, if the equipment is active and requires specific periodic services (maintenance), while the service centers are heterogeneous can only provide limited services. Also, if these centers are managed in an integrated manner, providing a suitable schedule for the repair and maintenance of the existing facilities/equipment is a necessity [3].

With the advent of recent technology for making good products with high quality and designing highly reliable and efficient systems, the importance of the maintenance activities and maintenance management has considerably risen in all sectors of manufacturing companies and service organizations [3, 5].

Many researchers and practitioners are interested to contribute in improving the equipment availability level, cost-effectiveness, performance efficiency, on-time delivery, product quality, and environmental requirements, and so on. [6-9]. In other words, maintenance strategy is applied to enhance the reliability and robustness of the equipment by reducing unplanned downtimes, eliminating unforeseen failures, and minimizing the maintenance costs that play a significant role in reaching reliability and safety requirements [10-12].

However, since the capacity of maintenance service centers is limited in each period, some services may not be available if these periodic services are not scheduled. Thus, an optimal schedule enhances the reliability of each facility reduces the costs of the company.

Owing to the flexibility and dynamic changes of the target market, the classic centralized production planning and scheduling methods and their mechanisms are no longer responsive. This makes the companies to establish their facilities and service centers in dispersed places to satisfy the customers. This problem can be well fitted into unrelated parallel machines category, but it is so complicated due to the dispersed production and service centers.

In some instances, the scheduling problems are more complex than the above-mentioned conditions and occasionally some customers with geographical dispersed locations request a product or service which can be satisfied by different suppliers, based on the product/service brand. Each of these suppliers has different service providers in different places. Therefore, the plan should identify the main supplier and then the correspondent servicer and finally specify the job sequence for the selected servicer. Accordingly, the MFP scheduling problems are NP-hard, which means that finding their solutions without explicit enumeration methods is impossible and their computational time exponentially increases as the size of the problem raises. Moreover,
determining the optimum solution using mixed integer linear programming (MILP) model is not efficient, especially in large-scale instances [13].

MILP model can merely solve small-scale problems of MFP scheduling problem, while meta-heuristic techniques are usually applied to handle large-scale ones. Here, heuristic methods can only solve the parallel machines with small scale while they lose their capability/efficiency as the problem size increases and closes to the real size [14].

As a real case study of an MFP system, consider the compressed natural gas (CNG) stations controlled by a supervisor, where there is a centralized supervision and management which controls the activities of all stations. The repair times are defined by maintenance experts based on preventive maintenance models and then the central management should implement the repair plan with the lowest cost and earliness/tardiness using the existing heterogeneous companies.

The scheduling of periodic maintenance services of heterogeneous multi-agent companies with limited capacity to customers in dispersed locations can be considered as an important problem of the MFP system. In some of the conducted research in MFP, the simple form of this problem has been studied; i.e., only a single product is delivered to a customer or set of customers, only one time period is considered, the factories are considered homogeneous, and other simple assumptions, which neglect the real-world conditions. However, the proposed model and solution approach in this study mostly stand on real assumptions.

The first contribution of this paper is to develop a novel bi-objective optimization model in which real assumptions, including heterogeneous manufacturers with limited capacity, multi-period service scheduling, soft time window in providing services, and geographically dispersed locations of factories and customers, are considered in the MFP system. The second contribution is to present both exact and meta-heuristics methods for solving the small-scale and large-scale problems, respectively. The third contribution is to tackle the uncertainty of the parameters using robust possibilistic programming. Moreover, this study includes a case study of Iranian CNG stations periodic maintenance service scheduling as a real problem.

In summary, the main purpose of this research is to provide a bi-objective model and solution approach to solve maintenance scheduling and planning in the heterogeneous MFP system. This model minimizes the total tardiness / earliness in the execution of the maintenance under the uncertainty. This research is organized into six sections. In Section 2, the literature is reviewed. In Section 3, the problem is described. In Section 4, the proposed bi-objective optimization model, robust programming approach for dealing with uncertainty, and the exact and meta-heuristics solution approaches are explained. In Section 5, the case study and numerical result analysis are presented. Section 6 concludes the paper with a discussion of the results and suggesting some directions for future research.
2. The Literature Review

A large variety of techniques have been developed for parallel machine scheduling (PMS). For example, Balakrishnan et al. [15] studied unrelated PMS using MILP model. They could successfully plan a two PMS problem with up to 10 jobs. In this area, Zhu et al. [16] also modelled an MILP for unrelated PMS.

An exact solution based on branch and bound (B&B) technique was used to solve unrelated PMS problem with 30 jobs [17]. Furthermore, other researchers, such as Ruiz [18], proposed some heuristic methods to solve smaller scale PMS problems and then evaluated their methods with the above-mentioned exact techniques.

Kanyalkar et al. [19] categorized MFP as unrelated PMS problems. They introduced MFP including its differences with single (centralized) factory production. The products are produced in multiple factories in MFP and these manufacturers may be positioned in dispersed locations. Therefore, some of these factories may be close to the customers, while others may not. Nevertheless, all factories are not capable to accomplish all jobs. In other words, there is something called “capability of factory” and the capacity of factories is different with each other.

Behnamian and Ghomi [1] considered a MFP model, where each factory had some parallel machines and each of which might have different speed. Thus, the jobs processing time could be different in each factory. The goal was to minimize the completion time or makespan. This problem could be investigated in unrelated PMS category, however it was assumed that the machines in the same group are identical, while each group has different machines from other groups. They also presented a GA for solving large-scale instances in addition to the design of a computational model for this problem.

A complicated study was performed on multi-factory scheduling with limited service using a real case study in jersey production factory in Belgium [20]. The objective was defined as the minimization of the weighted combination of delay and earliness. The due dates and change times are sequence dependent. Their research was developed considering the geographical dispersion of manufacturing sites.

A review of the multi-factory machine scheduling for the first time was provided by Behnamian and Fatemi Ghomi [21]. This paper classified and reviewed the literature in terms of shop environments, including single machine, parallel machines, flow shop, job shop, and open shop.

An unrelated parallel machine scheduling (UPMS) problem was considered with time-dependent deterioration and multiple rate-modifying activities by [22]. In the proposed UPMS problem, they simultaneously determined the schedule of the jobs and the number and positions of rate-modifying activities to minimize the makespan. In this paper, extensive computational experiments were also conducted through randomly generated examples to evaluate the performance of the proposed algorithms.

Mensendiek et al. [23] addressed the problem of minimizing the total tardiness of a set of jobs to be scheduled on identical parallel machines, where jobs could only be delivered at certain fixed delivery dates. The authors developed and empirically evaluated both the optimal and
heuristic solution procedures to solve their problem. The results proved that both approaches provided optimal solutions for instances with less than 20 jobs and different tightness of delivery dates in a reasonable computational time.

Poursabzi et al. [24] studied the problem of capacitated lot-sizing and scheduling in job shops with a carryover set-up and a general product structure. They first developed an efficient mixed integer linear programming (MILP) model for the problem, and then, they adapted an available lower bound (LB) in the literature to their problem. Some heuristic methods based on the production shifting concept were also proposed to solve this problem.

The scheduling problem in a hybrid flow shop (HFS) with unrelated parallel machines was investigated by [25]. In this paper, a Lagrangian Relaxation (LR) algorithm was developed to handle the HFS scheduling problem and two approaches were designed, namely, simplification of sub-problems and dominance rules, to solve the sub-problems generated in each iteration.

Furthermore, during the last few decades, numerous papers with various methods have been published on maintenance modeling and optimization [11, 12, 26-32]. For example, Garg et al. [27] presented the periodic preventive maintenance (PM) of a system with deteriorated components, in which PM simultaneously considered three action of mechanical service, repair and replacement for a multi-components system, based on maintenance cost. In this article, the degraded behavior of the component was modeled by a reliability equation, and the effect of PM actions to reliability was formulated based on the maximization of the maintenance-benefit analysis. They also presented a two-phase approach for the statistical analysis of failure data of a crank-case manufacturing of a two-wheeler industry, covering a period of one year. In this paper, for getting the global values of the parameters probability distribution of failure and repair, the particle swarm optimization was developed [28].

Niwas and Garg [3] presented an approach for analyzing the behavior of an industrial system under the cost-free warranty policy. The distribution of failure and repair time was assumed to be negative exponential, and various parameters such as reliability, mean time to system failure, availability and expected profit were derived for a system by using a mathematical modeling with Markov process. In another work, a greedy heuristic-based local search algorithm was developed to provide a system maintenance schedule for multi-component systems, coordinating the recommended component maintenance times to reduce system downtime costs. The minimization of the sum of downtime, earliness and tardiness costs of scheduling was defined as the objective of the proposed iterative algorithm [29].

Usually, in the works on scheduling, population-based meta-heuristics are used more frequently to solve these types of problems. Among them, it has been shown that the GAs usually had better performance than other population-based and local search algorithms [33]. For example, GA is developed for the optimization of the maintenance scheduling of generating units in a power system by [34]. An efficient GA was used to solve the resource-constrained project scheduling problem by [35]. An extended GA was proposed for solving open-shop scheduling problem by [36]. A hybrid GA approach was presented for preventive remanufacturing planning of production equipment under operational and imperfect maintenance constraints by [36].
addition, GA is still of great interest to researchers not only for scheduling but also for nonlinear constrained optimization problems [37].

Maintenance optimization is a multi-objective problem in nature, and it usually needs to achieve a trade-off between time and reliability objectives. The multi-objective meta-heuristic methods based on GA, for example non-dominated sorting genetic algorithm (NSGA-II), are applied for solving bi-objective scheduling problems. Recently, NSGA-II was used to solve the imperfect preventive maintenance optimization [38]. It handles the multi-objective optimization of parallel machine scheduling integrated with multi-resources preventive maintenance planning [38]. Moreover, it was employed for energy-efficient job shop scheduling [39]. Regarding the efficient performance of the GA and NSGA-II, we will use this approach to solve the large-scale scheduling problem that is described in detail in future sections.

3. Problem definition

In this paper, scheduling of periodic services to customers in dispersed locations is studied in which the customers have different services in each period. The service centers are also positioned in dispersed locations each of which has limited number of teams capable to offer one or some services. The duration time and cost of services are also considered heterogeneous.

Suppose a network including two levels/echelons. In one level, there are service centers with limited number of teams and facilities, while the customers are in other level as the applicants for services. Both levels are geographically dispersed in vast regions. The customers are placed in different places and there are one or some teams in each center capable of providing one or some services with different time and cost. The goal is to simultaneously minimize ‘total service costs’ and ‘total earliness/tardiness’ in providing services to customers through optimum assignment of required customers’ services to service centers and scheduling of offering services.

The services offering are periodically carried out. The required services by each customer is determined in each period and a soft time window is taken into account to fulfill such a need. Considering the offered services as “jobs”, existing teams in each factory as “machines” and time needed to provide each service as “processing time of machine, this problem can be stated as a Unrelated Parallel Machine Scheduling problem (UPMS) one to present services from multiple factories to multiple customers. The goal is to seek the optimal scheduling to provide services in each period to customers by existing teams in each factory, where tardiness/earliness and total costs are minimized. In this paper, Unrelated Parallel Machine Scheduling problem with Multiple Factories/server and Clients called UPMS_MFC is investigated. Table 1 shows the characteristics of this problem based on the notation in the literature.

[Please insert Table 1 about here]
The other assumptions in this study used in modeling and solving the problem are as follows:

- It is possible to perform a service before or after its due date, up to a certain limit determined as upper or lower bounds.
- No pre-emption of services is allowed, i.e., the services should be completely delivered/presented after start.
- Number of factories/service centers and service teams are finite.
- Number of client centers and their required services are limited.
- The planning horizon comprises finite number of time periods with given and fixed length.
- A given service/job may be ordered by a customer in several periods.
- All services should be performed according to the planning horizon, however some services belonging to a given period may be presented in successor or predecessor periods.
- Each team returns to its factory after accomplishing the service in each period, and then get prepared to do next service and going toward customer.
- The duration time of each service accomplished by each team is uncertain.

4. Modeling and solution approach

In this section, solution approach is described. First, a bi-objective mixed linear programming model along with augmented epsilon constraint to exactly solve the studied problem. Then, a bi-objective meta-heuristic algorithm based on GA is proposed and its performance is assessed to see whether one can apply it to solve large-scale instances or not. Uncertainty of some parameters such as service time is handled and controlled through robust possibilistic programming.

4.1. Bi-objective optimization model

Sets and indices

- \( F = \{1, 2, \ldots, F\} \) Index for factories/service centers
- \( K = \{1, 2, \ldots, K\} \) Index for customers/client centers
- \( J = \{1, 2, \ldots, J\} \) Index for jobs/services
- \( S_f = \{1, 2, \ldots, S_f\} \) Index for teams/machines in factory \( f \)
- \( H = \{1, 2, \ldots, H\} \) Index for time periods/days (e.g. months of a year)

Parameters/input data

- \( d_{ijk} \) 1 if time period \( t \) is predetermined due date to do service \( j \) for customer \( k \); 0 otherwise
- \( \bar{\varepsilon}_j \) Maximum deviation from due date of service \( j \) (based on time period). The \( \left[ (t - \bar{\varepsilon}_j) d_{ijk}, (t + \bar{\varepsilon}_j) d_{ijk} \right] \) deadline is considered for providing services.
### Decision variables/outputs

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{jkti}$</td>
<td>1 if service $j$ of customer $k$ in period $t$ is offered by team $i$ belongs to factory $f$; 0 otherwise</td>
</tr>
<tr>
<td>$y_{tjkti}$</td>
<td>1 if in period $t''$, service $j$ of customer $k$ in period $t$ is offered before service $j'$ of customer $k'$ in period $t'$ by team $i$ belongs to factory $f$; 0 otherwise</td>
</tr>
<tr>
<td>$z_{jkt}$</td>
<td>1 if service $j$ belongs to customer $k$ in period $t$ is fulfilled with tardiness; 0 otherwise</td>
</tr>
<tr>
<td>$c_{jkt}$</td>
<td>Duration time of service $j$ belongs to customer $k$ in period $t$ which is done in period $t'$</td>
</tr>
<tr>
<td>$T_{jkt}$</td>
<td>Tardiness in fulfilling service $j$ belongs to customer $k$ in period $t$</td>
</tr>
<tr>
<td>$E_{jkt}$</td>
<td>Earliness in fulfilling service $j$ belongs to customer $k$ in period $t$</td>
</tr>
</tbody>
</table>

### Objective Functions

**Minimize** SC: 

$$
SC = \sum_{j} \sum_{k} \sum_{t} (v_{jkt} + t_{c_{jkt}}) x_{jkti} + \sum_{j} \sum_{k} \sum_{t} f_{jkt} z_{jkt} 
$$

(1)

**Minimize** ET: 

$$
ET = \sum_{j} \sum_{k} \sum_{t} w_{j} (\omega T_{jkt} + (1-\omega) E_{jkt})
$$

(2)

Eq. (1) minimizes total service costs including operational costs of service, transportation cost of teams from service centers to clients’ cites and fixed cost of tardiness/earliness in providing services. Eq. (2) total weighted tardiness and earliness in presenting services are minimized. In Eq. (2), $\omega$ and $(1-\omega)$ are importance factors of having no tardiness or earliness, respectively. If both tardiness and earliness are equally important, the parameters $\omega$ is considered as $\omega \equiv 0.5$, but if earliness has no effect in providing services, this parameters is considered as $\omega \equiv 1$. 

The constraints of the studied problem are as follows:

\[
\sum_{t'} \sum_{j} \sum_{f} x_{jtk'i} = d_{ij} \quad \forall \, j \in J, t \in T, k \in K \tag{3}
\]

\[
\sum_{t} \sum_{i} \sum_{f} x_{jtk'i} \leq 1 \quad \forall \, j \in J, t' \in T, k \in K \tag{4}
\]

\[
\sum_{i} \sum_{j} \sum_{f} \sum_{k} x_{jtk'i} \leq 3 \quad \forall \, t' \in T, k \in K \tag{5}
\]

\[
\sum_{i} \sum_{j} \sum_{k} \sum_{t} \sum_{f} x_{jtk'i} \leq 2 \quad \forall \, i \in S_f, f \in F, t' \in T \tag{6}
\]

\[
x_{jtk'i} \leq \sum_{t'} \sum_{j'} \sum_{k'} y_{j't'k'i} \quad \forall \, j \in J, t,t'' \in T, k \in K, i \in S_f, f \in F \tag{7}
\]

\[
\sum_{j} \sum_{k} y_{jtk'i} \leq 1 \quad \forall \, t,t'' \in T, i \in S_f, f \in F \tag{8}
\]

\[
y_{jtk'i} + y_{jtk'i} \leq 1 \quad \forall \, t,t' \in T, i \in S_f, f \in F, j, j' \in J, k, k' \in K \tag{9}
\]

\[
y_{jtk'i} + y_{jtk'i} \geq 1 \quad \forall \, t,t',t'' \in T, i \in S_f, f \in F, j, j', j'' \in J, k, k', k'' \in K \tag{10}
\]

\[
c_{jkt} - c_{jkt} + S_{ijfk} + p_{j} + (y_{jtk'i} - 2)M \quad \forall \, t,t' \in T, k, k' \in K, j, j' \in J, i \in S_f, f \in F \tag{11}
\]

\[
z_{jtk} = d_{ij} \sum_{t'} \sum_{j} \sum_{f} x_{jtk'i} \quad \forall \, j \in J, t \in T, k \in K \tag{13}
\]

\[
T_{jtk} = d_{ij} \left[ \left( \sum_{t'} \sum_{i} \sum_{f} t'x_{jtk'i} - t - 1 \right) + \frac{c_{jkt}}{\tau} \right] \quad \sum_{t'} \sum_{i} \sum_{f} t'x_{jtk'i} - t > 0 \tag{14}
\]

\[
E_{jtk} = d_{ij} \left[ \left( t - \sum_{i} \sum_{f} \sum_{t'} t'x_{jtk'i} \right) - \frac{c_{jkt}}{\tau} \right] \quad \sum_{t'} \sum_{i} \sum_{f} t'x_{jtk'i} - t < 0 \tag{14}
\]

\[
\forall \, j \in J, t \in T, k \in K \tag{15}
\]

\[
l_{jtk} \geq \sum_{t'} \sum_{i} \sum_{f} x_{jtk'i} \quad \forall \, j \in J, t \in T, k \in K \tag{15}
\]
\[ l_{jtk} \leq 1 - \sum_{r'} \sum_{t} \sum_{f} x_{jtkr'f} \quad \forall j \in J, t \in T, k \in K \] (16)

\[ T_{jtk} \geq d_{jtk} \left[ \left( \sum_{r'} \sum_{t} \sum_{f} t' x_{jtkr'f} - t - 1 \right) + \sum_{r'} \frac{c_{jtkr'}}{T} \right] + (l_{jtk} - 1)M ; \forall j \in J, t \in T, k \in K \] (17)

\[ E_{jtk} \geq d_{jtk} \left[ \left( t - \sum_{r'} \sum_{t} \sum_{f} t' x_{jtkr'f} \right) - \sum_{r'} \frac{c_{jtkr'}}{T} \right] - (l_{jtk} + (1 - z_{jtk}))M ; \] (18)

\[ \forall j \in J, t \in T, k \in K \]

\[ T_{jtk} \leq \xi_{j} + M \left( 1 - d_{jtk} \right) \forall j \in J, t \in T, k \in K \] (19)

\[ E_{jtk} \leq \xi_{j} + M \left( 1 - d_{jtk} \right) \forall j \in J, t \in T, k \in K \] (20)

\[ c_{jtkr'} \leq \tau ; \forall j \in J, t, t' \in T, k \in K \] (21)

\[ \left\{ \begin{array}{l} x_{jtkr'f}, y_{jtkr'f}, z_{jtkr}, l_{jtk} \in \{0,1\} \\ c_{jtkr'}, T_{jtk}, E_{jtk} \geq 0 \end{array} \right. \] (22)

Eq. (3) ensures that each required service should be offered for each customer in each period. Constraint (4) ensures that each customer can only request a given service at most once in each period. Eq. (5) controls maximum number of services belonging to each customer in each period, e.g., each customer can use three different services in each period.

Eq. (6) shows maximum number of services for each team, e.g., each team of each factory can present at most two different services. Eq. (7) shows that if any team services to a customer, it means either this is the first provided service by that team ( \( y_{1j} = 1 \) ), or this team has already provided another service ( \( y_{0j} = 0 \)).

It is obvious that the first service offered by each team (if any) in each period is unique. This constraint is satisfied through Eqs. (8-9) show the precedence of two consecutive different services. It should be mentioned that no team can provide two different services simultaneously. It is also obvious that the sequence of providing services by each team in each period has transitivity property shown in Eq. (10).

In Eq. (11), duration time of a service is calculated, only if another service was already presented. According to this constraint, if a given team is willing to do specific service ( \( x_{jtk} = 1 \) ), while another service was already done by that team ( \( y_{1j} = 1 \) ), the completion time of the second service equals to completion time of the first service plus setup time for the second service (which equals to total needed time of returning the team from first customer to factory
plus needed time to transport the team to the second customer, i.e., \( S_{ij'k} = t_{kj} + t_{jk} \) as well as processing time of the second service. It should be mentioned that this constraint will be dummy if any of mentioned prerequisites is not active \( (x_{jk'f'i'g} = 0) \) and Eq. (12) will be active which calculates duration time of each team’s first service.

Eq. (13) determines which services have either tardiness or earliness. If a given service belongs to period \( t (d_{jk} = 1) \) and is presented in any other period except period \( t \) \( \sum \sum \sum \sum x_{jkif} = 1 \), it can be concluded that this service is done with time deviation from its due date. It should also be mentioned that if period \( t \) hasn’t already been assigned to a given service \( (d_{jk} = 0) \), this equation will be dummy and consequently \( z_{jk} = 0 \).

Eq (14) calculates tardiness and earliness in providing required services of each customer. This constraint can be linearized using Eqs. (15-16). To do so, the auxiliary binary variable \( l_{jk} \) should be first defined as follows: 1, if a given service is presented with tardiness \( \sum \sum \sum \sum x_{jkif} - t > 0 \), 0 otherwise. Then, using this variable and \( z_{jk} \) which was already explained, it is obvious that either tardiness will be added to the second objective function or earliness or none of them.

Eqs. (17-20) limit the maximum allowable tardiness and earliness in providing services to an upper bound. Eq. (21) controls the service time duration in each period. Finally, Eq. (22) shows decision variables and their domain including some binary variables and some nonnegative ones.

4.1.1. Uncertainty control using robust possibilistic programming approach

Usually there is no complete certainty in most of parameters in real optimization problems, while obvious uncertainty is companied with them. The problem solution may be inefficient if these uncertainties are not controlled. In studied UPMS_MFC in this paper, the processing time of each job by each machine (duration time of presented service by each team, i.e., \( p_{jk'f} \)) is considered uncertain. In this regard, mathematical programming techniques are applied as follows to handle such uncertainty, since this uncertain parameter can be stated in terms of a fuzzy number.

Suppose \( T \) is a passive parameter. Although, one cannot exactly determine its value, however it is possible to limit it into a given interval of numbers according to previous knowledge, experience and expert estimation, where this value with different probabilities is equal to any of existing numbers of this interval. For example, consider four numbers \( t_1 < t_2 \leq t_3 < t_4 \), where \( T \) cannot take value lesser than \( t_1 \) or more than \( t_4 \) (such probability is negligible). Also, the most probability belongs to values between \( t_2 \) and \( t_3 \). The probability of
being equal to any number between \( t_1 \) and \( t_2 \) is linearly increasing, while the probability of being equal to any number between \( t_3 \) and \( t_4 \) is linearly decreasing. Based on the said, \( \tilde{T} \) is limited to a trapezoid set so called trapezoidal fuzzy number (TFN) depicted in Fig. 1. To handle the uncertain parameter \( \tilde{p}_{jif} \) in studied UPMS_MFS in this research, it is assumed that it could be stated according to the experts’ opinions and historical data as follows.

If some parameters are TFN in an optimization problem, one can employ different approaches such as possibilistic programming to solve such a problem [40, 41] as a subset of fuzzy mathematical programming/optimization [42]. In the following, possibilistic programming is first briefly described and then its newfangled expression integrated with robustness concept is presented. Finally, the robust version of UPMS_MFC is employed in this research.

Consider the following fuzzy programming model:

\[
\begin{align*}
\text{Min} & \ cx \\
\text{s.t.} & \ Ax \geq \tilde{b} \\
& \ x \in X
\end{align*}
\]  

(23)

Where \( \tilde{b} = (b^1, b^2, b^3, b^4) \) is a TFN vector and \( b^1 < b^2 \leq b^3 < b^4 \) are vectors of crisp numbers. According to possibilistic programming, a level \( (\alpha \times 100)\% \) is first taken into account for constraints and then a possibility measure \( \alpha \) for constraints is considered as follows:

\[
\begin{align*}
\text{Min} & \ cx \\
\text{s.t.} & \ Poss \left( Ax \geq \tilde{b} \right) \geq \alpha \\
& \ x \in X
\end{align*}
\]  

(24)

According to the possibility measure, the above-mentioned possibilistic programming equals to the following defuzzified model [40, 41]:

\[
\begin{align*}
\text{Min} & \ cx \\
\text{s.t.} & \ Ax \geq \alpha b^4 + (1-\alpha) b^3 \\
& \ x \in X
\end{align*}
\]  

(25)

To improve the latter model performance, in conducted research by Pishvae et al. [43] by considering two concepts, i.e., feasibility robustness and optimality robustness, robust
possibilistic programming (RPP) were developed, where possibility measure $\alpha$ were interactively
determined according to robustness concept within the problem solution process. The general
form of RPP is as follows:

$$
\begin{align*}
\text{Minimize } & \phantom{\alpha} \text{cx} + \phi \left[ b^4 - (\alpha b^4 + (1-\alpha) b^3) \right] \\
\text{subject to } & \phantom{\alpha} Ax \geq \alpha b^4 + (1-\alpha) b^3 \\
& \phantom{\alpha} x \in X \\
& \phantom{\alpha} 0.5 \leq \alpha \leq 1
\end{align*}
$$

(26)

Where $\phi \geq 0$ is a control parameter obtained by sensitivity analysis. Also, $\alpha$ is a variable
signifying that how much the constraints can be applied. It is obvious that if $\phi \to 0$, then $\alpha \to 0.5$ and if $\phi \to \infty$, then $\alpha \to 1$. According to the above-mentioned explanations, in RPP
approach in solving of defined UPMS_MFC problem in this study, the parameter

$$
P_{ijf} = \left( P_{ijf}^1, P_{ijf}^2, P_{ijf}^3, P_{ijf}^4 \right)
$$

is first stated as a TFN and then constraint (27) is replaced with

$$
c_{jkt'} = r_{jkt'} + \sum_{k} s_{i,j;k,k} y_{q,i;jk} \geq \left( \alpha P_{ijf}^4 + (1-\alpha) P_{ijf}^3 \right) x_{jkt';ijf} ;
$$

(27)

$$
\forall j \in J, t, t' \in T, k \in K, i \in S_j, f \in F
$$

Now, constraints (28) and (29) in calculating of objective functions should be rewritten in
the same way as follows:

$$
\text{Minimize } SC = \sum_{j} \sum_{t} \sum_{k} \sum_{t'} \sum_{i} \sum_{f} \left[ v_{jktf} + tc_{jkt} \right] x_{jkt'if} + \sum_{i} \sum_{f} f_{jkt} z_{jkt} + \\
\phi_1 \sum_{j} \sum_{i} \sum_{f} \left[ P_{ijf}^4 - (\alpha P_{ijf}^4 + (1-\alpha) P_{ijf}^3) \right] 
$$

(28)

$$
\text{Minimize } ET = \sum_{j} \sum_{i} \sum_{f} w_{j} \left( \omega T_{jkt} + (1-\omega) E_{jkt} \right) + \\
\phi_2 \sum_{j} \sum_{i} \sum_{f} \left[ P_{ijf}^4 - (\alpha P_{ijf}^4 + (1-\alpha) P_{ijf}^3) \right] 
$$

(29)

4.2. Trade-off between objectives using augmented epsilon constraint method

As already pointed out, the goals of solving UPMS_MFC is satisfying two objectives
simultaneously, i.e., minimizing total costs of providing services to customers and minimizing
total tardiness and earliness. In practice, there is contradiction between objectives, namely, by
increasing the quality of one objective, the quality of another one decreases and vice versa. Accordingly, different approaches are proposed to solve bi- or multi-objective decision making (MODM) problems such as weighted sum method (WSM), epsilon constraint (EC), augmented epsilon constraint (AEC), goal programming (GP), lexicographic (Lex), etc [44, 45].

The general form of a MODM problem, Eq. (30), is as follows:

\[
\begin{align*}
\min & \left( f_1(x), f_2(x), \ldots, f_n(x) \right) \\
\text{s.t.} & \quad x \in X
\end{align*}
\] (30)

In EC method, one objective is first considered as main objective, while the rest of objectives are limited to upper bound of epsilon \( e_i \) and the following single-objective model is obtained in terms of Eq. (31):

\[
\begin{align*}
\min & \quad f_i(x) \\
\text{s.t.} & \quad f_i(x) \leq e_i, i = 2, 3, \ldots, n \\
& \quad x \in X
\end{align*}
\] (31)

In EC method, by changing the values of \( e_i \), different solutions are obtained which may not be efficient (weakly efficient). This difficulty has been rectified in AEC method through replacing following model [46]:

\[
\begin{align*}
\min & \quad f_i(x) - \sum_{i=2}^{n} \phi_i s_i \\
\text{s.t.} & \quad f_i(x) + s_i = e_i, i = 2, 3, \ldots, n \\
& \quad x \in X \\
& \quad s_i \geq 0
\end{align*}
\] (32)

Where the \( s_i \) are nonnegative variables for shortage and \( \phi_i \) is a parameter used for normalizing the first objective function’s value with respect to \( i^{th} \) objective \( \phi_i = R(f_i) / R(f_i) \). To better implement the AEC method, one can obtain the appropriate interval of epsilons \( e_s \) using Lex method [47].

In order to apply AEC method in solving of defined UPMS_MFC, the first objective function (minimizing total service costs, i.e., \( f_1 = SC \)) is considered as the main objective, while the second objective (minimizing total weighted tardiness and earliness, i.e., \( f_2 = ET \)) is limited to different values of epsilons and efficient solutions can be then obtained using Eq. (32) for different values of \( e \).
4.2. Meta-heuristic solution approach

As already mentioned, complexity of PMS problems and large sizes of real instances make the exact methods based on mathematical programming (such as proposed MILP) inefficient. The exact mathematical models can only solve this type of problems in small sizes, while meta-heuristic algorithms are usually used to solve large-scale cases [48-51] Accordingly, and since the defined UPMS_MFC in this study is a more complex version of PMS, besides the proposed mathematical model in previous section, a bi-objective meta-heuristic method based on GA, i.e., non-dominated sorting genetic algorithm II called NSGA-II is applied to solve the large-scale instances efficiently.

The most important parts of NSGA-II are determining structure of the chromosomes (solution representation, neighborhood structure, crossover and mutation) and fitness function. The structure of chromosomes should be included at least all model’s variables as well as most of the problem constraints. On the other hand, crossover and mutation operators should be easily applied on this structure. The structure of crossover and mutation should be defined in a way that the solution space can be completely explored with the capability of generating high-quality solutions. The main parts of NSGA-II are precisely explained as follows.

4.2.1. Chromosome structure

The designed chromosome for the studied problem in this research comprises two rows. The number of columns is equal to the number of orders by customers from all services in all periods calculated based on parameter $d_{ijk}$. In other words, number of columns in this matrix is equal to those number of elements in $d_{ijk}$ matrix which take value. In the following, it will be determined that each customer’s order from each service can be fulfilled by which teams. Suppose $t=3$, $j=3$, $k=2$ and $S=4$, so a sample of this matrix is as follows:

Please insert Figure 2 about here

In Fig. 2, a sample of chromosome structure designed for the studied problem in this paper is presented. In the first row, $d_{1,2,2}$ signifies demand/order of customer 2 for service 2 in period 1, and $d_{1,1,1}$ indicates order of customer 1 for service 1 in period 1. In the second row and first column, it can be observed that the order of customer 2 for service 2 in the first period is fulfilled by team #2. It should be mentioned that in generating of chromosome’s row, each order/demand is determined by those teams having capability of providing that service. For instance, if teams #1, #2 and #4 can satisfy demand of $d_{1,2,2}$, one of them is randomly selected and is lied in the second row of the chromosome.

So, the first row signifies the sequence of providing services to customers for each service in each period, while the second row determines each demand/order (specified in the first row) can
be fulfilled by which team. In other words, the presented chromosome structure shows both sequence of providing services to customers and assignment of demands/orders to teams.

4.2.2. Fitness function

After determining the chromosome structure, one should evaluate the objective function. According to Fig. 2 it can be observed that the demands $d_{1.2.2}$, $d_{1.2.1}$, $d_{3.1.2}$, and $d_{2.1.2}$ are fulfilled by team #2. According to the first row of chromosome, the sequence of providing services to customers by team #2 is determined. Team #2 goes from factory to customer 2 at first and presents service 2; then goes to customer 1 to offer service 2. Next, goes to customer 2 to present service 2 at first and subsequently service 1. According to this sequence, one can easily calculate the time of requested services by customers. On the other hand, the chromosome structure is designed in a way that any type of requested service by any customer is accomplished by those teams having the capability of doing it. This rule is applied to all teams. Finally, after determining the service time requested by each customer, one can calculate the tardiness or earliness in providing any demanded service by customers in each period. A penalty function is defined to limit the maximum possible amount of tardiness and earliness in presenting services. Suppose there is a constraint as follows, inequality (33):

$$C_i(x) \leq C$$  \hfill (33)

The penalty function for these constraints is defined as follows:

$$Vio = \max \left( \frac{C_i(x)}{C} - 1.0 \right)$$  \hfill (34)

Eq. (34) as penalty function is calculated for maximum amount of tardiness and earliness and added to the main objective functions as Eqs. (35) and (36) in terms of multiplier:

$$Z_1 = F_1(x) + \alpha_1 Vio$$  \hfill (35)

$$Z_2 = F_2(x) + \alpha_2 Vio$$  \hfill (36)

Where, $F_1(x)$ and $F_2(x)$ are the values of the first and second objective function and multipliers $\alpha_1$ and $\alpha_2$ determine amount of penalty effect over each of objective functions.

4.2.3. Crossover

The crossover operator increases the diversity/dispersion of solutions and investigates the solution space extensively. In the proposed algorithm, a single-point crossover operator is used to generate the offsprings. In this method, having chosen two parents to mate, a point is randomly selected in chromosome as cut point. Then, the right parts of cut points are
interchanged and consequently two new offsprings are generated. Applying this method, the generated children exploit their both parents’ information/characteristics. Fig. 3 depicts the applied crossover in this study (4 points are randomly chosen as cut points).

[Please insert Figure 3 about here]

4.2.4. Mutation

As already pointed out, mutation operator avoids zeroing the probability of exploring each point of solution space. In other words, regardless of other existing members in population, mutation operators applies small changes over chromosome so as to increase the quality of obtained good solutions during optimization process to a possible extent. In this mechanism, two mutation operators are applied: swap and reversion.

In this type of mutation, two columns of chromosome are first randomly selected and their values are interchanged. Fig. 4 depicts a given sample of this type.

[Please insert Figure 4 about here]

As it can be observed in Fig. 4, columns 4 and 8 are randomly selected and their positions are then swapped. For instance, in the initial chromosome, team #4 presents service 1 to customer 1 and then goes to customer 2 to fulfill service 3. The occurred change through mutation operator causes that team #4 goes first to customer 2 to present service 3 and then refers to customer 1 to accomplish service 1. The similar changes are happened to team #2.

In this type of mutation, two columns of initial chromosome are first selected randomly and the columns between these selected columns are reversed from right to left. Fig. 5 illustrates a given sample of this type.

[Please insert Figure 5 about here]

According to Fig. 5, columns 4 and 8 are selected as mutation points, and then columns 4 to 8 are reversed from right to left and the new offsprings are then obtained.

4.2.5. Stopping criterion:

Among different stopping criteria proposed in the literature, reaching to a predetermined number of iterations/generations is set as stop criterion in the applied NSGA-II in this study.

5. Analysis and evaluation of results
In this section, proposed model and solution approaches and also their applicability to solve real problems are evaluated. To do so, maintenance scheduling of CNG stations equipment in Iran as a real case study in defined UPMS_MFC area is investigated in this research. According to this case study, a small-sized sample is first defined and outputs of different solution approaches are then analyzed and evaluated. Next, some representations of this case in different sizes (number of CNG stations, number of repairs in different periods, number of factories and servicing teams, etc. change the problem size) are presented and proposed solution approaches are evaluated and compared. The mathematical model is coded in GAMS 24.7.1 and solved by the CPLEX solver on a PC with a 2.5 GHz Intel® Core™ i5 processor and 6 GB RAM memory. Also, NSGA-II is coded in MATLAB 2016b.

5.1. Case study
As already mentioned, the investigated case in this research studies UPMS_MFC problem to yield the optimum maintenance scheduling of CNG stations equipment in Iran. These stations are usually located in different places, each of which has specific equipment (such as compressor, dispenser, dryer, etc.) which have supplied from different brands and companies as depicted in Fig. 6. Each of these equipment has usually standard norm for control and repair. The supervisors of these stations offer the existing fundamental equipment along with their brand and forecasted maintenance scheduling in each station to the general manager of all CNG stations.

National Iranian Oil Products Distribution Company (NIOPDC) is responsible for integrated management of all CNG stations in Iran. Moreover, implementing the optimum maintenance scheduling of CNG stations are in charge of NIOPDC and its managers try to save expenditures and yield minimum possible amount of tardiness and earliness through running such an optimum schedule.

Existing equipment in each CNG station are purchased from different brands and companies which these factories are located in in different places. Suppose equipment E belongs to brand B. If this equipment is maintained by instruments of brand B (its supplier), the minimum time and cost should be spent, otherwise this repair should be carried out by other brands imposing more time and cost.

[Please insert Figure 6 about here]

In each factory, there is finite number of teams to implement customers maintenance plan (CNG stations in this study) each of which fulfills specific maintenance (not necessarily any type of maintenance). In addition to provided information of CNG stations (forecasted maintenance scheduling), related statistics about establishment places of factories and their limitation in offering services along with maintenance expenditures are reported to managers of NIOPDC.
It is clear that NIOPDC could implement the maintenance scheduling without any tardiness or earliness and with minimum cost, if number of servicing teams in each factory are infinite with thorough availability to all regions. But, in practice limitation of servicing teams from one side and dispersion of CNG stations from other side (may be caused inaccessibility to some factories due to long distance) result in much complexity in running this schedule and making decision will be difficult about it. This obstacle becomes more unintelligible when the maintenance schedule should be covered more number of stations, more types of repairs, and more number of factories and servicing teams. The investigated case study, is a sample of defined UPMS_MFC problem in this research which could be solved using proposed model and solution approaches, where the results can significantly help the managers of NIOPDC to make the best decisions.
5.2. Validation of proposed solution approaches

In this subsection, a representation of explored case study in small size (with 13 stations, 5 factories with 13 servicing teams, 5 types of service correspondent to those equipment in Fig. 6) is presented in Fig. 7, where the initial evaluation of proposed solution approaches are carried out using this small sample. Table 2 shows maintenance scheduling of CNG stations (annually) and it can be observed that each station needs to which maintenance in which periods.

[Please insert Figure 7 about here]

[Please insert Table 2 about here]

5.2.1. Evaluation of AEC method compared to NSGA-II

In order to obtain the optimal/global Pareto front, the proposed mathematical model and AEC exact method are first employed for solving the studied bi-objective problem in small sizes. The NSGA-II method is then employed for the same reason and its obtained Pareto front is compared with the Pareto front gained from hybrid method LexAEC (hybridization of AEC with Lex method). The first objective function is determined with ‘Cost’, while the second one is specified with time window unsatisfaction (TWU) in Tables 3 and 4. Fig. 8 and Fig. 9 show the Pareto fronts of AEC method and NSGA-II respectively and they are compared by Fig. (10).

In order to compare the results of AEC exact method and NSGA-II, a small instance of UPMS_MFC problem is solved. The obtained Pareto fronts of both algorithms are depicted in Fig. 10, simultaneously. Since this problem is small sized, it was already also anticipated that the AEC method can outperform NSGA-II, however NSGA-II has acceptable performance in this sample instance, where its Pareto front is close to global front gained by AEC method to a large extent. In practice, a solution should be opted from Pareto optimal front by managers/decision makers through doing trade-off between solutions. In Fig. 11, a given space is suggested to select the Pareto front among different obtained solutions, since the rate of costs increment is more than rate of costs decrease to a large extent.

[Please insert Table 3 about here]

[Please insert Table 4 about here]

[Please insert Figure 8 about here]

[Please insert Figure 9 about here]

[Please insert Figure 10 about here]

[Please insert Figure 11 about here]

In order to compare the results of AEC exact method and NSGA-II, a small instance of UPMS_MFC problem is solved. The obtained Pareto fronts of both algorithms are depicted in
Fig. 10, simultaneously. Since this problem is small sized, it was already also anticipated that the AEC method can outperform NSGA-II, however NSGA-II has acceptable performance in this sample instance, where its Pareto front is close to global front gained by AEC method to a large extent. In practice, a solution should be opted from Pareto optimal front by managers/decision makers through doing trade-off between solutions. In Fig. 11, a given space is suggested to select the Pareto front among different obtained solutions, since the rate of costs increment is more than rate of costs decrease to a large extent.

5.2.2. Evaluation of RPP approach to control uncertainty

In this subsection, RPP approach is assessed in control of UPMS_MFC’s uncertainty defined in this research. The two criteria “deviation from optimality” and “constraints violation” are from the most important indices for performance evaluating of optimization approaches in uncertain conditions. To use these criteria, the uncertain parameter in this study, i.e., processing times (duration time of maintenance in case study) are simulated 20 times and the performance of proposed RPP approach is evaluated.

It is assumed that the average of fuzzy data \( \frac{P_{ij}^1 + 2P_{ij}^2 + 2P_{ij}^3 + P_{ij}^4}{6} \) in nominal value approach is replaced with them. In robust possibilistic (Robust I) approach this value is already determined similar to possibility measure \( \alpha \) (\( \alpha = 95\% \) is considered in this research). Finally, in RPP approach (Robust II) is \( \alpha \) considered a variable obtained through solving of the model. According to the gained results in Figs. 12 and 13, it can be observed that optimality fluctuations in proposed robust approaches is very lesser than nominal value approach. Secondly, the proposed robust possibilistic approaches significantly shrink constraints violations than nominal value approach, which itself diminishes the risk of decision making.

[Please insert Figure 12 about here]

[Please insert Figure 13 about here]

5.3. Evaluation of solution approaches for large-size problems

In subsection 5.2.1, a representation of UPMS_MFC problem was solved in small size, where according to its results, it could be observed that the Pareto fronts obtained from NSGA-II was to a large extent close to global optimal Pareto fronts illustrating its acceptable performance. In this subsection, validation of the proposed meta-heuristic algorithm, NSGA-II, is evaluated in a more comprehensive space for large-size instances. To do so, some measures are first defined and some experimental instances in different sizes are then designed. Results of AEC exact method and NSGA-II as a meta-heuristic algorithm is compared in Tables 5-8.
5.3.1. Cover set (CS)

In this criterion, the number of non-dominated solutions in each method is compared with other method’s ones [51]. Consider two solution approaches A and B for a given MODM problem, where \( F(A) \) and \( F(B) \) signify Pareto fronts obtained from solution approach A and B, respectively. Also, for each member \( pa \in F(A) \) and \( pb \in F(B) \), the symbol \( pa \ Dom pb \) indicates dominancy of \( pa \) against \( pb \) (or equivalently being dominated \( pb \) against \( pa \)).

According to these definitions, the measure \( CS \) is introduced so as to compare the two solution approaches A and B as follows:

\[
CS(A, B) = \left\{ pb \in F(B) \mid \exists pa \in F(A) : pa \ Dom pb \right\} / F(B)
\]

As a matter of fact, the measure \( CS(A, B) \) shows the portion of total Pareto solutions of method B which are dominated by at least one of the Pareto solutions of method A. It is obvious that \( 0 \leq CS(A, B) \leq 1 \).

- If \( CS(A, B) \) is close to 0, then method B has better performance than A and most of its solutions are efficient.
- If \( CS(A, B) \) is close to 1, then method A has better performance than B and most of its solutions are efficient.
- The less the value of \( CS(A, B) \), the better the performance of method B.

5.3.2. Mean of ideal distance (MID):

In this criterion, as one of the most important criteria for measurement of MODM problems [34], an ideal solution is first considered for the on-hand problem and mean deviations of Pareto solutions from ideal solutions are then calculated. The ideal solution shown by \( I_{sol} \) is called to an status in which both solutions are simultaneously optimum, i.e., \( I_{sol} = (\min(Z_1), \min(Z_2)) \). It is obvious that in problems in which all of objective functions are “minimization”, one can set the origin of the coordinate as the ideal solutions, i.e., \( I_{sol} = (0,0) \).

If \( F(A) \) signifies Pareto front obtained from solution approach A, MID criterion is calculated as follows:

\[
MID(A) = \frac{\sum_{pa \in F(A)} |pa - I_{sol}|^2}{F(A)}
\]

Where, \( I_{sol} - pa \) shows the Euclidean distance of solutions \( pa \in F(A) \) from ideal solutions. Clearly, the less the value of MID criterion, the better its performance.
5.3.3. Number of solutions (NOS) or solutions quantity:

In this criterion, the number of obtained Pareto solutions are computed. The method with more number of solutions (NOS) is better. For method A, this criterion is stated as \( NOS(A) = F(A) \). Despite usefulness of NOS for measuring the diversity of solutions, however it has a major weakness; the quality of solutions cannot be clearly observed. This obstacle is rectified in next measure.

5.3.4. Number of non-dominated solutions (NS_CS) or solutions quality:

One of the weaknesses of NOS is when \( NOS(B) > NOS(A) \), while \( CS(A,B) \) is a large number. This means that the most of obtained solutions by method B are dominated by those gained from method A. However, according to NOS, as it can be seen, method B outperforms method A. To fix this difficulty, a hybrid criterion called NS_CS is introduced as follows:

\[
NS\_CS(A,B) = \left[ NOS(B)(1-CS(A,B)) \right]
\]  

(39)

In fact, \( NS\_CS(A,B) \), counts the number of Pareto solutions obtained from method B which are not dominated by solutions of method A. It is evident that the more the value of \( NS\_CS(A,B) \), the better the performance of method B.

[Please insert Table 5 about here]

[Please insert Table 6 about here]

[Please insert Table 7 about here]

[Please insert Table 8 about here]

6. Conclusions and future studies

In this research, the scheduling of periodic services from heterogeneous multi-agent companies to customers located in dispersed locations and have different needs and services is investigated. For this problem, named as UPMS_MFC, two objective functions are considered: service costs and tardiness/earliness minimization. To solve this problem, first, a bi-objective mixed integer linear programming (MILP) model which is handled by augmented epsilon constraint (AEC) is developed, and then, a meta-heuristic method named as NSGA-II is proposed. In addition, to handle the uncertainty of some parameters, the robust possibilistic programming (RPP) approach is employed.

To evaluate the performance of proposed bi-objective MILP and NSGA-II solution methods, several experimental problems have been randomly generated and different criteria such as MID, NOS and NS_CS were used. The obtained results showed that the global Pareto
fronts could be gained for small size instances using the proposed AEC exact method. Also, NSGA-II had comparable performance against AEC in small-sized instances, which is acceptable. This guarantees that one can employ NSGA-II in large-size problems for which AEC is not capable to solve the problems. Furthermore, to show the stability of the proposed meta-heuristics method in solving large-scale test problems, the NSGA-II method has been implemented several times for each experimental problem. According to the results, the performance is acceptable and the proposed NSGA-II approach is reliable for solving various large-scale problems.

Using simulated numerical instances, it could be observed that ‘constraints violation’ and ‘deviation from optimality’, as two important indices of the optimization approaches performance in uncertain conditions, significantly decrease in RPP approach, which in turn, diminish the risk of decision making.

As a stream for future studies, one can consider the impact of maintenance scheduling on the reliability of equipment in a company, wherein the probability of activity interference for each facility is less than a predetermined bound. Another interesting direction can be to take the uncertainty of the other parameters into account and present different powerful meta-heuristic algorithms to tackle the studied problem.

References


Hamed Jafar-Zanjani is currently a PhD candidate in Industrial Engineering with specialty in automation at Science and Research Branch, Islamic Azad University, Tehran, Iran. He received his BSc degree in Industrial Engineering from Naragh Branch, Islamic Azad University, Markazi, Iran, in 2003. And his MSc degree in Industrial Engineering from Najaf Abad Branch, Islamic Azad University, Isfahan, Iran, in 2006. He has more than thirteen years of experience in planning and project control in industrial. His current research interests are scheduling and optimization techniques.

Mostafa Zandieh received his BSc degree in Industrial Engineering from Amirkabir University of Technology, Tehran, Iran, in 1998, and MSc degree in Industrial Engineering from Sharif University of Technology, Tehran, Iran, in 2000. He also obtained his PhD degree in Industrial Engineering from Amirkabir University of Technology, Tehran, Iran, in 2006. Currently, he is an Associate Professor in Industrial Management Department at Shahid Beheshti University, Tehran, Iran. His research interests are production planning and scheduling, financial engineering, quality engineering, applied operations research, simulation, and artificial intelligence techniques in the areas of manufacturing systems design.

Mohammad Khalilzadeh received his MSc and PhD degrees in Industrial Engineering from Sharif University of Technology, Tehran, Iran. He is currently an Assistant Professor in the Industrial Engineering Department of the Science and Research Branch of the Islamic Azad University, Tehran, Iran. His research interests are in project management, fuzzy multiple-criteria decision making, and system thinking.
### Tables

**Table 1. The characteristics of studied UPMS_MFC problem in this research**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Abbreviation symbol</th>
<th>Explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Machines</strong> (servicer teams) [α]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UR: Unrelated machines</td>
<td></td>
<td>Each factory or service center has multiple machines/teams with different speed and cost, which can process in parallel.</td>
</tr>
<tr>
<td>DD: Due date</td>
<td></td>
<td>A due date is considered for ordered services.</td>
</tr>
<tr>
<td>RD: Release date</td>
<td></td>
<td>Start time of each job/service depends on its availability. In other words, all jobs cannot be presented at the outset of planning horizon. It is also possible that all ordered services in each period cannot be presented at the outset of that period.</td>
</tr>
<tr>
<td>SD: Sequence dependent setup time</td>
<td></td>
<td>The setup time of each machine/team depends on the jobs sequence. This time significantly depends on distance between different client centers and their correspondent service centers.</td>
</tr>
<tr>
<td>ER: Eligibility restriction</td>
<td></td>
<td>Each factory’s teams cannot offer all services. In other words, there is a limitation in proficiency of machines/teams.</td>
</tr>
<tr>
<td>N_PC: No precedence constraints</td>
<td></td>
<td>There is no precedence or posterior in providing services, i.e., no service is predecessor or successor of another one (no job should be done before or after that job).</td>
</tr>
<tr>
<td>N_B: No breakdown</td>
<td></td>
<td>All machines/service teams are permanently available, i.e., no breakdown is required.</td>
</tr>
<tr>
<td>Objective function (desired to solve the problem) [γ]</td>
<td>N_BP: No batch processing</td>
<td>SC: Service cost</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>----------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td></td>
<td>Each machine/team can only perform one process or service at the time, i.e., there is no way to do some services by a team.</td>
<td>Total service costs should be minimized including operational costs of service by each team for each service in each service center, transportation cost of teams for carrying them from service centers to clients’ cites, fixed cost of tardiness/earliness in providing services.</td>
</tr>
<tr>
<td>ET: Earliness and tardiness</td>
<td>Total weighted of tardiness and earliness in providing services</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Maintenance scheduling of CNG stations (annually)

<table>
<thead>
<tr>
<th>CNG Stations</th>
<th>Equipment and maintenance periods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Compressor</td>
</tr>
<tr>
<td>CNG1</td>
<td>9-5-1</td>
</tr>
<tr>
<td>CNG2</td>
<td>9-6-2</td>
</tr>
<tr>
<td>CNG3</td>
<td>9-5-1</td>
</tr>
<tr>
<td>CNG4</td>
<td>9-5-1</td>
</tr>
<tr>
<td>CNG5</td>
<td>10-6-2</td>
</tr>
<tr>
<td>CNG6</td>
<td>9-5-2</td>
</tr>
<tr>
<td>CNG7</td>
<td>10-5-1</td>
</tr>
<tr>
<td>CNG8</td>
<td>10-5-1</td>
</tr>
<tr>
<td>CNG9</td>
<td>10-5-1</td>
</tr>
<tr>
<td>CNG10</td>
<td>10-5-1</td>
</tr>
<tr>
<td>CNG11</td>
<td>10-5-1</td>
</tr>
<tr>
<td>CNG12</td>
<td>10-5-1</td>
</tr>
<tr>
<td>CNG13</td>
<td>10-6-2</td>
</tr>
</tbody>
</table>
Table 3. Trade-off between cost and customers’ unsatisfaction using AEC method

<table>
<thead>
<tr>
<th>Pareto solution</th>
<th>First objective function’s value (Cost)</th>
<th>Second objective function’s value (TWU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7127</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6513</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>6110</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>5510</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>5178</td>
<td>40</td>
</tr>
<tr>
<td>6</td>
<td>4650</td>
<td>60</td>
</tr>
<tr>
<td>7</td>
<td>4045</td>
<td>130</td>
</tr>
<tr>
<td>8</td>
<td>3750</td>
<td>190</td>
</tr>
<tr>
<td>9</td>
<td>3625</td>
<td>230</td>
</tr>
<tr>
<td>10</td>
<td>3310</td>
<td>280</td>
</tr>
<tr>
<td>11</td>
<td>3150</td>
<td>340</td>
</tr>
<tr>
<td>12</td>
<td>3098</td>
<td>370</td>
</tr>
</tbody>
</table>
Table 4. Trade-off between cost and customers’ unsatisfaction using NSGA-II method

<table>
<thead>
<tr>
<th>Pareto solution</th>
<th>First objective function’s value (Cost)</th>
<th>Second objective function’s value (TWU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7349</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6513</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>5890</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>5178</td>
<td>40</td>
</tr>
<tr>
<td>5</td>
<td>4850</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>4245</td>
<td>140</td>
</tr>
<tr>
<td>7</td>
<td>3690</td>
<td>180</td>
</tr>
<tr>
<td>8</td>
<td>3625</td>
<td>230</td>
</tr>
<tr>
<td>9</td>
<td>3512</td>
<td>280</td>
</tr>
<tr>
<td>10</td>
<td>3150</td>
<td>340</td>
</tr>
<tr>
<td>11</td>
<td>3098</td>
<td>370</td>
</tr>
<tr>
<td>12</td>
<td>7349</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 5. Scale of UPMS_MFC test problems (small-scaled)

<table>
<thead>
<tr>
<th>No. of instance</th>
<th>Planning periods</th>
<th>maintenance/ job</th>
<th>Locations / CNG stations</th>
<th>Teams</th>
<th>Factory</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>2</td>
<td>7</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>3</td>
<td>10</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>4</td>
<td>10</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>4</td>
<td>12</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>5</td>
<td>14</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>5</td>
<td>15</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>6</td>
<td>20</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>
Table 6. Scale of UPMS_MFC test problems (large-scale)

<table>
<thead>
<tr>
<th>No. of instance</th>
<th>Planning periods</th>
<th>maintenance/job</th>
<th>Locations / CNG stations</th>
<th>Teams</th>
<th>Factory</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>5</td>
<td>30</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>6</td>
<td>30</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>7</td>
<td>30</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>8</td>
<td>30</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>9</td>
<td>35</td>
<td>13</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>10</td>
<td>40</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>10</td>
<td>45</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>10</td>
<td>50</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>12</td>
<td>11</td>
<td>60</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>12</td>
<td>70</td>
<td>22</td>
<td>20</td>
</tr>
</tbody>
</table>
Table 7. Comparison of proposed solution approaches according to evaluation measures (small-scaled)

<table>
<thead>
<tr>
<th>instance</th>
<th>CS (AEC, NSGAII)</th>
<th>MID (AEC)</th>
<th>MID (NSGAII)</th>
<th>NOS (AEC)</th>
<th>NOS (NSGAII)</th>
<th>NS_CS (AEC, NSGAII)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>150.83</td>
<td>150.83</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>134.60</td>
<td>134.60</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>0.33</td>
<td>221.89</td>
<td>203.12</td>
<td>5</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>304.05</td>
<td>287.24</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>0.20</td>
<td>287.51</td>
<td>275.31</td>
<td>11</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>351.43</td>
<td>390.64</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>0.14</td>
<td>400.85</td>
<td>430.65</td>
<td>16</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>0.07</td>
<td>531.15</td>
<td>494.65</td>
<td>17</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>560.13</td>
<td>559.08</td>
<td>17</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>587.42</td>
<td>604.15</td>
<td>19</td>
<td>17</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 8. Performance of proposed NSGA-II and its stability in large-scale instances

<table>
<thead>
<tr>
<th>instance</th>
<th>MID</th>
<th>NOS</th>
<th>Run Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>B</td>
<td>W</td>
</tr>
<tr>
<td>1</td>
<td>820.65</td>
<td>812.23</td>
<td>903.21</td>
</tr>
<tr>
<td>2</td>
<td>928.76</td>
<td>873.3</td>
<td>980.43</td>
</tr>
<tr>
<td>3</td>
<td>1070.67</td>
<td>1008.37</td>
<td>1090.43</td>
</tr>
<tr>
<td>4</td>
<td>1324.59</td>
<td>1279.49</td>
<td>1333.56</td>
</tr>
<tr>
<td>5</td>
<td>1351.73</td>
<td>1343.73</td>
<td>1411.73</td>
</tr>
<tr>
<td>6</td>
<td>1377.90</td>
<td>1289.43</td>
<td>1448.65</td>
</tr>
<tr>
<td>7</td>
<td>1630.59</td>
<td>1572.43</td>
<td>1697.65</td>
</tr>
<tr>
<td>8</td>
<td>1635.80</td>
<td>1578.21</td>
<td>1728.54</td>
</tr>
<tr>
<td>9</td>
<td>1746.24</td>
<td>1665.24</td>
<td>1766.51</td>
</tr>
<tr>
<td>10</td>
<td>1875.79</td>
<td>1868.73</td>
<td>1933.53</td>
</tr>
</tbody>
</table>

Abbreviation: M: mean; B: best; W: worst; SD/M: standard deviation per mean
Figures

Fig. 1. A trapezoidal fuzzy number

Fig. 2. A sample of chromosome structure

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>$d_{122}$</th>
<th>$d_{111}$</th>
<th>$d_{232}$</th>
<th>$d_{121}$</th>
<th>$d_{312}$</th>
<th>$d_{112}$</th>
<th>$d_{311}$</th>
<th>$d_{132}$</th>
<th>$d_{212}$</th>
<th>$d_{322}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
### Fig. 3. Crossover operator

<table>
<thead>
<tr>
<th>Parent 1</th>
<th>$d_{1,1,1}$</th>
<th>$d_{1,2,2}$</th>
<th>$d_{2,1,1}$</th>
<th>$d_{2,2,1}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,3,2}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,2,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parent 2</th>
<th>$d_{1,1,1}$</th>
<th>$d_{1,2,2}$</th>
<th>$d_{2,1,1}$</th>
<th>$d_{2,2,1}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,3,2}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,2,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Child 1</th>
<th>$d_{1,1,1}$</th>
<th>$d_{1,2,2}$</th>
<th>$d_{2,1,1}$</th>
<th>$d_{2,2,1}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,3,2}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,2,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Child 2</th>
<th>$d_{1,1,1}$</th>
<th>$d_{1,2,2}$</th>
<th>$d_{2,1,1}$</th>
<th>$d_{2,2,1}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,3,2}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,2,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

### Fig. 4. A given sample of swap mutation

<table>
<thead>
<tr>
<th>Parent</th>
<th>$d_{1,1,1}$</th>
<th>$d_{1,2,2}$</th>
<th>$d_{2,1,1}$</th>
<th>$d_{2,2,1}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,3,2}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,2,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Child</th>
<th>$d_{1,1,1}$</th>
<th>$d_{1,2,2}$</th>
<th>$d_{2,1,1}$</th>
<th>$d_{2,2,1}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,3,2}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,2,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

### Fig. 5. A given sample of reversion mutation

<table>
<thead>
<tr>
<th>Parent</th>
<th>$d_{1,1,1}$</th>
<th>$d_{1,2,2}$</th>
<th>$d_{2,1,1}$</th>
<th>$d_{2,2,1}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,3,2}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,2,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Child</th>
<th>$d_{1,1,1}$</th>
<th>$d_{1,2,2}$</th>
<th>$d_{2,1,1}$</th>
<th>$d_{2,2,1}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,3,2}$</th>
<th>$d_{3,1,2}$</th>
<th>$d_{3,2,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Fig. 6. Some of existing fundamental equipment in a CNG station requiring to periodic maintenance

Fig. 7. A representation of providing maintenance services network from factories to CNG stations
Fig. 8. Pareto front obtained by AEC method

Fig. 9. Pareto front obtained by NSGA-II method
Fig. 10. Comparison of Pareto fronts obtained by AEC and NSGA-II methods

Fig. 11. Suggested area to select Pareto solutions
Fig. 12. “Deviation from optimality” criteria in comparison of robust and nominal approaches

Fig. 13. “Constraints violation” criteria in comparison of robust and nominal approaches