A genetic algorithm for subcontractors selection and allocation in multiple building projects

Mohammad Reza Afshar1 · Vahid Shahhosseini1 · Mohammad Hassan Sebt1

Accepted: 28 June 2021 / Published online: 19 July 2021 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
In construction industry, general contractors (GCs) need to manage and conduct numerous projects simultaneously. To this aim, they usually have subcontractors conduct available tasks and projects. So, subcontractor management is becoming a major challenge. To deal with this challenge, GCs aim to reduce total costs of projects including employment/subcontracting costs, indirect costs, tardiness penalties, and the money which must be paid for movement of workforce from one project to another. The aim of current research study is to propose a model for selecting subcontractors and assigning available tasks in the project to them in order to reduce the costs of the GC. Then, a genetic algorithm is proposed to solve a real problem. The proposed algorithm is innovative from three points of view: (1) generation of initial population; (2) subcontractor assignment approach; and (3) the objective function. The problem was also solved by means of an exact method. Then, the results of proposed algorithm were compared to the outcomes of the exact method. This comparison shows that the proposed algorithm can efficiently help GCs select subcontractors and assign available tasks to each subcontractor when several projects must be carried out simultaneously.

Keywords Subcontractors selection · Subcontractors assignment · Genetic algorithm · Multiple project environment

1 Introduction

One of the initial steps in project planning is to identify the right subcontractors and assign the available work packages of a project to them. In this research study, the word subcontractor means the contractors who are working as some groups at the lowest level in the hierarchy of contractors in a project. For example, the ones who are involved in concreting or reinforcing operations.

Pells (1993) mentioned that “the selection of subcontractors is a difficult and time-consuming process.” General contractors (GCs) usually have limited information about the real cost and duration of a project because these parameters depend on the capabilities of the subcontractors who carry out the work packages of the project.

Many researchers have studied subcontractor selection in the literature to introduce quantitative approaches to help GCs in this regard (Monghasemi et al. 2015). Most of these approaches have been based on multi-criteria decision-making approaches (Afshar et al. 2020a, b; Afshar et al. 2017; Fachruzazi and Munirwansyah 2017; Ulubeyli and Kazaz 2016; Polat 2016). Recent research studies show that a significant proportion of them did not consider the impact of subcontractor selection on overall features of the project including project duration, cost, and feasibility. While these features are closely linked to the performance of subcontractors. Hence, mathematical models were applied more to consider the effect of subcontractor selection on the overall features of a project (Afshar et al. 2020a, b; Sonmez and Gürel 2016; Polat et al. 2015; Bešikçi et al. 2015; Mungle et al. 2013; Mokhtari and Abadi 2013; Polack-Johnson and Liberatore 2006; Józefowska et al. 2001). However, most of these models have studied subcontractor selection when the GC had to conduct only one project.
Whereas, most of GCs must carry out several projects simultaneously (El-Abbasy et al. 2017). To make the problem more realistic, it is necessary to consider the condition in which the GCs need to conduct numerous projects at the same time. Taking these realistic situations in consideration makes the model and the formulation more complicated. To deal with these realistic problems, Afshar et al. (2020a, b) investigated the problem of subcontractor selection when the GC must conduct numerous projects at the same time. However, their model did not consider the appropriate assumption of the projects in the construction industry. Additionally, more actions need to be taken in order to consider some other aspects of the project including the cost of employed skilled workers and subcontractors and the time which employed skilled workers and subcontractors need to move from one project to another one. So, it is necessary to reach a more practical approach.

Another drawback is that Afshar et al. (2020a, b) used the CPLEX solver of the general algebraic modeling system (GAMS) for solving the problem. However, the application of precise techniques would fail to solve the problem in a reasonable computing time because the problem itself was a non-deterministic polynomial-time hard (NP-hard) type and real construction projects include multiple work packages and subcontractors (Afshar et al. 2020a, b). Thus, in practice, metaheuristic algorithms are required for large construction projects to generate near-optimal solutions in reasonable computational time.

In order to fill the mentioned gaps, the current study tries to propose a real model for subcontractor selection and allocation in multiple building projects that fits the real situation. In this regard, first the required assumption will be discussed and then the model will be proposed. In this model, different locations, times, and effort which is required for the movement of employed skilled workers and subcontractors from one project to another are considered. Then, a GA is developed to solve the problem. Innovation of the proposed algorithm can be discussed in three ways:

This algorithm proposes a heuristic method to generate a part of the initial population. An infeasible tackling procedure is also presented to convert infeasible solutions to feasible ones, and a simple fitness function is proposed to handle feasible and infeasible solutions. To validate the efficiency and feasibility of the proposed model and proposed GA, a real case study of the building industry is included.

The remainder of the paper is organized as follows. Section 2 presents the problem description. The proposed procedure for solving the subcontractor selection and allocation problem is stated in the third section. The computational experiments are performed in Sect. 4 in order to prove the effectiveness of the proposed model and metaheuristic algorithm. In Sect. 5, managerial implications are discussed. In the end, some conclusions are presented in Sect. 6.

## 2 Problem statement

The list of notations that are used in this research study is presented in this part.

### Sets and indices

- \(i, h\) Set of work packages
- \(j, g\) Set of projects
- \(k\) Set of SCs
- \(f\) Set of discount levels
- \(SC_{ij}\) Set of SCs that can perform work package \(i\) of project \(j\)
- \(n_j\) Number of work packages of project \(j\)
- \(pw_{ij}\) Predecessors set of project \(j\)
- \(TC\) Total cost

### Parameters

- \(IC_j\) The indirect cost of project \(j\) per time period
- \(Cy_k\) Full-time employment of skilled workers that work for subcontractor \(k\) (the group of skilled workers \(k\)) per time period
- \(CT_{jg}\) Transfer cost of one subcontractor or one group of skilled workers from project \(j/g\) to project \(g/j\)
- \(Bp_{ijk}\) The bid price for work package \(i\) of project \(j\) proposed by SC \(k\)
- \(B_j\) Early completion bonus of project \(j\) as per earliness period
- \(l\) Maximum number of work packages that GC intends to sublet to each SC
- \(Sp_j\) The earliest start time of project \(j\)
- \(d_{ijk}\) Duration of work package \(i\) of project \(j\), operating by SC \(k\)
- \(d'_{ijk}\) Duration of work package \(i\) of project \(j\), operating by group of skilled workers \(k\)
- \(d_{jg}\) Transfer time of one subcontractor or one group of skilled workers from project \(j/g\) to project \(g/j\)
- \(L_{hij}\) The time lag between work package \(h\) and work package \(i\) in project \(j\)
- \(A_{fk}\) Minimum number of work packages that should be sublet to the SC \(k\) in order to activate the level \(f\) of his/her discount levels
- \(B_{fk}\) Maximum number of work packages that should be sublet to the SC \(k\) in order to activate the level \(f\) of his/her discount levels
- \(IF\) Incentive fee
- \(DP_{fk}\) Discount percentage of level \(f\) presented by SC \(k\)
- \(M\) A very large number
Due date of project $j$

Decision variables

- $T_j$: Completion time of project $j$ (integer variable)
- $T_{0k}$: Completion time of collaboration between GC and group of skilled workers $k$ (integer variable)
- $W_k$: Time of collaboration between GC and group of skilled workers $k$
- $N_k$: Percent of SC $k$'s bid prices paid to him/her when one of his/her discount levels is activated (positive variable)
- $n_{xijk}$: Linearization auxiliary variable (positive variable)
- $SW_{ijk}$: Start time of work package $i$ of project $j$, operating by SC $k$ (integer variable)
- $SSW_{ijk}$: Linearization auxiliary variable (integer variable)
- $SW_{0ij}$: Start time of work package $i$ of project $j$, operating by group of skilled workers $k$ (integer variable)
- $SSW_{0ijk}$: Linearization auxiliary variable (integer variable)
- $K_j$: The early completion period of project $j$ (integer variable)
- $G_j$: Tardy period of project $j$ (integer variable)

$x_{ijk} = \begin{cases} 
1 & \text{if work package } i \text{ of project } j \text{ is sublet to SC } k \\
0 & \text{otherwise}
\end{cases}$

$y_{ijk} = \begin{cases} 
1 & \text{if work package } i \text{ of project } j \text{ is operated by group of skilled workers } k \\
0 & \text{otherwise}
\end{cases}$

$U_k = \begin{cases} 
1 & \text{if group of skilled workers } k \text{ is employed on a full time basis by GC} \\
0 & \text{otherwise}
\end{cases}$

$P_{ijk} = \begin{cases} 
1 & \text{if subcontractor } k \text{ or group of skilled workers } k \text{ performs work package } i \text{ of project } j \text{ before work package } h \text{ of project } g \\
0 & \text{otherwise}
\end{cases}$

$E_{fk} = \begin{cases} 
1 & \text{if discount level } f \text{ is activated for SC } k \\
0 & \text{otherwise}
\end{cases}$

In order to propose a model for subcontractor selection when numerous projects must be done simultaneously, the first 12 assumptions of the manuscript will be discussed and then the model will be produced.

**Assumption 1** In order to prevent any possible controversy or interference between subcontractors or the groups of skilled workers, each work package is assigned to one subcontractor/one group of skilled workers. This assignment will be done through employing subcontractor $k$ or full-time employment of the group of skilled workers which work for subcontractor $k$ (From this point, as far as the end of the manuscript group of skilled workers $k$ means a group of skilled workers who works for subcontractor $k$ and which is employed on a full-time basis by GC to carry out the related work packages).

Equation 1 which is mentioned below guarantees this issue.

$$
\sum_{k=1}^{SC} y_{ijk} + \sum_{k=1}^{SC} x_{ijk} = 1 \quad \forall j, i \in n_j
$$

**Assumption 2** In order to restrict the influence of subcontractors on overall process of the project and to restrict their capability to stop the project, the number of work packages that can be assigned to each subcontractor will be limited. The level of limitations will be determined by the GC. Equation 2 ensures this policy.

$$
\sum_{j=1}^{N} \sum_{i=1}^{n_j} x_{ijk} \leq I \times (1 - U_k) \quad \forall k
$$

**Assumption 3** When the group of skilled workers $k$ is employed to carry out the projects, the GC has more capability to control them. So, no limitation will be set on the number of the work packages which the group of skilled workers $k$ can conduct. Equation 3 guarantees this situation.

$$
\sum_{j=1}^{N} \sum_{i=1}^{n_j} y_{ijk} \leq M \times U_k \quad \forall k
$$

**Assumption 4** When more than one work package is assigned to one subcontractor or one group of skilled workers, they might not be able to carry out the work packages simultaneously due to their limited capability. Equations 4 and 5 confirm this approach.
\[ P_{ijhgk} + P_{hgjk} \leq 1 \quad \forall i,j,h,g,k \quad (4) \]

\[
\left( SW'_{ijk} + d'_{ijk} \right)y_{ijk} + \left( SW_{ijk} + d_{ijk} \right)x_{ijk} \\
\leq y_{hgk} \times SW_{hgk} + x_{hgk} \times SW_{hgk} + M(1 - P_{ijhgk}) \quad i \neq h \quad (5)
\]

**Assumption 5** The movement of one subcontractor or one group of skilled workers from one work package to another in one project has been considered in assumption 4. However, the time and money which must be spent on these items have not been taken into account. This fact happens when the movement occurs among the work packages of different projects. Accordingly, the time which is required for subcontractors or the group of workers to move from one project to another one must be considered. Equations 6 and 7 show these issues. The cost of movement is also expressed in the objective function.

\[ P_{ijhgk} + P_{hgjk} \leq 1 \quad \forall i,j,h,g,k \quad (6) \]

\[
\left( SW'_{ijk} + d'_{ijk} \right)y_{ijk} + \left( SW_{ijk} + d_{ijk} \right)x_{ijk} + d'_g \leq y_{hgk} \times SW_{hgk} + x_{hgk} \times SW_{hgk} + M(1 - P_{ijhgk}) \quad j \neq g \quad (7)
\]

**Assumption 6** Since most of the precedence relations among the activities in conventional building projects are based on the finish to start criteria with lag (Sonmez and Gürel 2016), in this project the precedence relations are considered based on the above-mentioned criteria which is expressed in Equation 8.

\[
\left( SW'_{hgk} + d'_{hgk} \right)y_{hgk} + \sum_{k=1}^{sc} \left( SW_{hgk} + d_{hgk} \right)x_{hgk} + L_{wij} \\
\leq \sum_{k=1}^{sc} y_{ijk} \times SW_{ijk} + \sum_{k=1}^{sc} x_{ijk} \times SW_{ijk} \sum_{k=1}^{sc} \quad \forall j, (h, i) \in ap_{ij} \quad h \neq 0 \quad (8)
\]

**Assumption 7** Starting time of each project must be after the official date of that project which is mentioned in the contract. Equation 9 confirms this point.

\[ Sp_j \leq \sum_{k=1}^{sc} y_{ijk} \times SW'_{ijk} + \sum_{k=1}^{sc} x_{ijk} \times SW_{ijk} \quad \forall i,j \quad (9) \]

**Assumption 8** When the group of workers \( k \) is employed by the GC, the duration of their collaboration with the GC is calculated based on the deadline of the projects that they took part in. So, the duration of collaboration is calculated based on Eqs. 10. Equation 11 calculates the time that the group of skilled workers collaborates with the GC.

\[ T_{0h} - y_{ijk} \times SW'_{ijk} - M \times (1 - y_{ijk}) \leq W_k \quad \forall i,j,k \quad (10) \]

\[ T_{0h} \geq T_j - M \times (1 - y_{ijk}) \quad \forall i,j,k \quad (11) \]

**Assumption 9** If more than one work package is assigned to one subcontractor, the subcontractor would offer a discount because its unemployment costs will reduce under this circumstance. In this research project, the discount is defined by different levels of discount.

Each discount level includes a discount percentage which is applied to the subcontractor’s bid price. By increasing the number of work packages that are assigned to a subcontractor through subcontracting, the discount percentage will increase. Equations 12 and 13 activate one of the discount levels for subcontractor \( k \). Here, Eq. 14 calculates the percentage of subcontractor \( k \) bid price which must be paid to this subcontractor \( (N_k) \).

\[ A_{jk} - M(1 - E_{jk}) \leq \sum_{f=1}^{n} \sum_{i=1}^{n} x_{ijk} \leq B_{jk} + M(1 - E_{jk}) \quad \forall f,k \quad (12) \]

\[ \sum_{f=1}^{n} E_{jk} \leq 1 \quad \forall k \quad (13) \]

\[ 1 - \sum_{f=1}^{n} D_{jk} \times E_{jk} = N_k \quad \forall k \quad (14) \]

**Assumption 10** In order to improve the reputation of the GC in front of the main employer, all the projects must be finished within the specified deadline. Equation 15 ensures this point and Eq. 16 calculated the duration of the project \( j \).

\[ T_j \leq DD_j \quad \forall j \quad (15) \]

\[ \sum_{k=1}^{sc} \left( SW'_{ijk} + d'_{ijk} \right)y_{ijk} + \sum_{k=1}^{sc} \left( SW_{ijk} + d_{ijk} \right)x_{ijk} \leq T_j \quad \forall i,j \quad (16) \]

**Assumption 11** If the GC finishes the project sooner than the specified deadline, it will be rewarded. This point is calculated by Eq. 17.

\[ K_j = DD_j - T_j \quad \forall j \quad (17) \]
Assumption 12 Indirect costs are directly linked to the duration of each project and are determined based on the size of the project in objective function (Eq. 18). Determination of these costs is not the subject of study in the current research project. In this research project, it is assumed that these costs are determined based on the experience of the GC.

Under the mentioned constraints, the GC as an economic organization seeks a way to minimize the total cost. Thus, the objective function (Eq. 18) is equal to the total cost of the projects including the following terms:

a) The cost of employing group of skilled workers (k)

\[ \left( \sum_{k=1}^{K} C_{yk} \times W_{k} \right) \]

b) Payment to the subcontractors

\[ \left( \sum_{j=1}^{N} \sum_{n_{i}=1}^{n_{i}} \sum_{k=1}^{K} B_{p_{ijk}} \times N_{k} \times x_{ijk} \right) \]

c) Indirect costs

\[ \left( \sum_{j=1}^{N} IC_{j} \times T_{j} \right) \]

d) Incentive fee

\[ \left( \sum_{j=1}^{N} B_{jG_{j}} \right) \]

e) The money which GC should pay to subcontractors or the groups of employed skilled workers for their movement from one project to another one

\[ \left( \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{h=1}^{H} \sum_{g=1}^{G} \sum_{k=1}^{K} CT_{jg} P_{ijghk} \right) \]

\[ \text{MIN} \ TC = \sum_{k=1}^{K} C_{yk} \times W_{k} + \sum_{j=1}^{N} \sum_{n_{i}=1}^{n_{i}} \sum_{k=1}^{K} B_{p_{ijk}} \times N_{k} \times x_{ijk} + \sum_{j=1}^{N} IC_{j} \times T_{j} - \sum_{j=1}^{N} B_{jG_{j}} + \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{h=1}^{H} \sum_{g=1}^{G} \sum_{k=1}^{K} CT_{jg} P_{ijghk} \]

(18)

It is clear that the proposed model is nonlinear. Solving such a problem will be time-consuming even in small-scale problems. So, this nonlinear model will change to a linear model in the next chapters.

2.1 Linearization of the proposed model

When some decision variables are multiplied to each other \((x_{ijk} \times SW_{ijk}, y_{ijk} \times SW'_{ijk}, x_{ijk} \times N_{k})\), nonlinearity will be produced in the model. In order to avoid nonlinearity, auxiliary variables of \(SW_{ijk}/SW'_{ijk}/nx_{ijk}\) are replaced by the mentioned multiply operations \(\left( x_{ijk} \times SW_{ijk}, y_{ijk} \times SW'_{ijk}, x_{ijk} \times N_{k} \right) \). Plus, the relationship between auxiliary variables and the main variables (Eqs. 19 to 27) are defined in a way that the result becomes equal to the result of the multiply operations which were omitted.

\[ \text{SSW}_{ijk} \leq \text{SW}_{ijk} \quad \forall i, j, k \]  
\[ \text{SSW}_{ijk} \geq \text{SW}_{ijk} - M \times (1 - y_{ijk}) \quad \forall i, j, k \]  
\[ \text{SSW}_{ijk} \leq M \times y_{ijk} \quad \forall i, j, k \]  
\[ \text{SSW}_{ijk} \leq \text{SW}_{ijk} \quad \forall i, j, k \]  
\[ \text{SSW}_{ijk} \geq \text{SW}_{ijk} - M \times (1 - x_{ijk}) \quad \forall i, j, k \]  
\[ \text{SSW}_{ijk} \leq M \times x_{ijk} \quad \forall i, j, k \]  
\[ nx_{ijk} \leq N_{k} \quad \forall i, j, k \]  
\[ nx_{ijk} \geq N_{k} - M \times (1 - x_{ijk}) \quad \forall i, j, k \]  
\[ nx_{ijk} \leq M \times x_{ijk} \quad \forall i, j, k \]  

2.2 Genetic algorithm (GA) for solving the subcontractor selection and assignment problem

The problem of subcontractor selection and assignment of the available work packages to them can be considered a part of an NP-hard-type problem (Afshar et al. 2020a, b). Accordingly, exact methods cannot solve these problems in a reasonable amount of time (Rehman et al. 2020; Khan et al. 2020).

Therefore, GA, as a metaheuristic algorithm, is introduced to be used for solving this problem. GA is generally utilized to generate high-quality solutions for optimization problems by mutation, crossover, and selection operators (Agrawal et al. 2021).

3.1 Solution representation

Before a metaheuristic algorithm started to search, a suitable representation scheme must be determined. For this aim, in the current study, the random key representation (RKR) is employed to determine the work packages’ priorities for scheduling. Also, the subcontractor list representation is used for assigning the subcontractors to the
work packages. The subcontractor list assigns the work package \( i \) to subcontractor \( k \) and it identifies whether the schedule is feasible (considering Eq. 2) or not. For the representation of the employment way pertaining to all subcontractors, the binary representation is applied (when a subcontractor is not selected, its employment method is not considered).

Hence, three independent chromosomes \((\alpha, \beta, \gamma)\) are considered. \( \alpha \) and \( \beta \) constitute the first part of each solution, while \( \gamma \) represents the second part of it. The number of genes in chromosomes \( \alpha \) and \( \beta \) were set to be the same as the total number of the work packages that the GC must plan to do. Additionally, the number of genes in chromosome \( \gamma \) was set to be the same as the total number of subcontractors who made their bid to execute available work packages.

### 3.2 Primary population

There are two methods for population updating: (1) random method and (2) using a heuristic method. Each of the methods has advantages and disadvantages. The random method increases exploration, while decreases exploitation. On the contrary, using a heuristic method increases exploitation while decreases exploration. In the current study, in order to use advantages of the both methods, 50% of the initial population was generated by using a heuristic
method, and other population was produced by random method.

On the other hand, using random generation approach for producing all initial populations may violate the deadline limitation for all of them. To avoid this violation, 50% of the initial population was generated by using a heuristic method. The proposed heuristic (Fig. 1) that is applied on chromosome $b$ tries to distribute the work packages among all subcontractors in order to prevent the delay in conducting work packages.

In other words, if more than one work package is assigned to one subcontractor, it is disabled to perform them simultaneously due to its limited capacity. Therefore, one of the work packages will be delayed and the probability of exceeding its project duration will increase. The proposed method can prevent this issue.

### 3.3 Infeasible tackling procedure

Some solutions may violate Eq. 2. These solutions will change to feasible solutions by use of infeasible tackling procedure. Accordingly, one work package which had been assigned to a subcontractor who violated restriction number 2 is considered and it will be assigned to another subcontractor.

This process will be continued until the violation of Eq. 2 is removed. Afterward, the obtained feasible solutions are allowed to be entered into the fitness function assessment.

### 3.4 Fitness function

To evaluate the population, the solutions should be decoded with the help of schedule generation schemes (SGS). SGS is capable of converting a solution into a schedule. A fitness value is accordingly calculated for each solution. Some solutions may violate Eq. 15.

In order to tackle the infeasible solutions, two approaches can be adopted. The first one is to omit the infeasible solutions and the other one is to penalize the available infeasible ones (Lee and El-Sharkawi 2008). Since there might be a solution among the infeasible solution with a low level of impossibility, it is not reasonable to omit them all. So, in this research study the second approach will be adopted, and to this aim fitness function (Eq. 28) is considered.

$$
f(p) = \begin{cases} 
\text{total cost}, & \text{if } G_j \leq 0 \\
\text{total cost} + \text{very large cost} + \text{very large cost} \times G_j, & \text{otherwise}
\end{cases}
$$

$$G_j = T_j - DD_j \quad \forall j$$

When it comes to the feasibility of the solutions ($G_j \leq 0$), the fitness function will be equal to the project’s total cost. In the case of the infeasibility of the solution, the fitness function is equal to the total project cost in addition to the fixed and variable penalty.

In fact, two extra costs have been considered for the infeasible cost. The role of the fixed penalty is to reduce the
chance of infeasible solutions to produce offspring and control the population. Besides, the variable penalty is designed to separate good and bad infeasible solutions.

### 3.5 Population updating mechanism

After trial and error between the possible crossover operations, two points crossover has been selected. To this aim, the chances \( \frac{2}{3} \) and \( \frac{1}{3} \) are assigned for parts 1 and 2, respectively (Fig. 2).

Then, mutation shall be applied randomly only on one part of each solution (Fig. 3).

In order to produce evolved population, three populations of the previous generation (the population which was produced by crossover operator and the population which was produced by mutation operator) have been mixed up.

Then the most appropriate individuals with \( POP\_size \) have been moved to the next generation.

### 3.6 Stopping criterion

In scheduling problems, generally, the termination condition is the limit of the number of generated schedules. The number of generated schedules is calculated by Eq. (30).

\[
\text{Generated schedules} = \frac{PO\_size}{C^3} + \left( \frac{PO\_size}{C^3} \cdot P_c + \frac{PO\_size}{C^3} \cdot P_m \right) \cdot \text{Iteration}
\]

(30)

In the current research, the number of generated schedules is chosen at 5000. Thus, based on Eq. (30), the iteration number of GA is determined.
| Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID |
|----|------|------|---------|----|------|------|---------|----|------|------|---------|
| 1  | 0.032| d_k1 | 4       | 5  | 6    | Sc4  | 0.03   | d_k4 | 60  | 75  | 90  | Sc7   | 0.002 | d_k7 | 5     | 8    | 11 |
| 2  | 0.048| d_k2 | 3       | 4  | 5    | Sc5  | 0.036  | d_k5 | 55  | 70  | 85  | Sc8   | 0.002 | d_k8 | 4     | 7    | 10 |
| 3  | 0.08 | d_k3 | 2       | 3  | 4    | Sc6  | 0.042  | d_k6 | 52  | 67  | 82  | Sc9   | 0.004 | d_k9 | 2     | 6    | 9  |

| Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID |
|----|------|------|---------|----|------|------|---------|----|------|------|---------|
| 10 | 0.006| d_k10| 28      | 42 | 56   | Sc13 | 0.013  | d_k13| 30  | 40  | 50  | Sc16  | 0.027 | d_k16| 20    | 25   | 30 |
| 11 | 0.007| d_k11| 26      | 40 | 53   | Sc14 | 0.014  | d_k14| 28  | 36  | 45  | Sc17  | 0.031 | d_k17| 18    | 22   | 25 |
| 12 | 0.008| d_k12| 24      | 37 | 50   | Sc15 | 0.015  | d_k15| 27  | 35  | 43  | Sc18  | 0.034 | d_k18| 17    | 21   | 24 |

| Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID |
|----|------|------|---------|----|------|------|---------|----|------|------|---------|
| 19 | 0.019| d_k19| 15      | 17 | 19   | Sc22 | 0.009  | d_k22| 30  | 45  | 60  | Sc25  | 0.008 | d_k25| 17    | 26   | 34 |
| 20 | 0.021| d_k20| 14      | 15 | 17   | Sc23 | 0.011  | d_k23| 28  | 42  | 55  | Sc26  | 0.009 | d_k26| 15    | 25   | 32 |
| 21 | 0.023| d_k21| 13      | 14 | 15   | Sc24 | 0.012  | d_k24| 27  | 40  | 53  | Sc27  | 0.013 | d_k27| 12    | 22   | 31 |

| Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID |
|----|------|------|---------|----|------|------|---------|----|------|------|---------|
| 28 | 0.003| d_k28| 8       | 12 | 16   | Sc31 | 0.036  | d_k31| 4   | 6   | 8   | Sc34  | 0.016 | d_k34| 10    | 15   | 20 |
| 29 | 0.003| d_k29| 7       | 10 | 14   | Sc32 | 0.053  | d_k32| 3   | 5   | 7   | Sc35  | 0.020 | d_k35| 9     | 14   | 19 |
| 30 | 0.004| d_k30| 6       | 9  | 11   | Sc33 | 0.056  | d_k33| 3   | 4   | 6   | Sc36  | 0.024 | d_k36| 8     | 13   | 18 |

| Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID | Sc | Cy-k | Item | Task ID |
|----|------|------|---------|----|------|------|---------|----|------|------|---------|
| 37 | 0.019| d_k37| 3       | 5  | 7    | Sc40 | 0.011  | d_k40| 5   | 8   | 12  | Sc43  | 0.002 | d_k43| 5     | 8    | 11 |
| 38 | 0.032| d_k38| 2       | 4  | 6    | Sc41 | 0.016  | d_k41| 4   | 7   | 10  | Sc44  | 0.002 | d_k44| 4     | 7    | 10 |
| 39 | 0.064| d_k39| 1       | 2  | 3    | Sc42 | 0.024  | d_k42| 3   | 6   | 8   | Sc45  | 0.003 | d_k45| 3     | 6    | 9  |
Table 3 Contractual and other data of projects which are related to the case study [duration time (days) and cost (1000$)]

| Data type                  | Item                  | Project |
|----------------------------|-----------------------|---------|
| Contractual data           | Project start time (day) [$S_{pj}$] | 0       |
|                           | Due date of project $j$ ($D_{Dj}$) | 220     |
|                           | Early completion bonus of project $j$ ($B_{j}$) | 0.01    |
| Other data                 | $I$                   | 3       |
|                           | Indirect cost of project $j$ per day ($I_{Cj}$) | 0.001   |

Table 4 Discount levels and related discount percentage proposed by the subcontractors

Table 5 Transfer time and costs of subcontractors among the projects

4 Computational experiments

The performance of the proposed mixed-integer linear programming (MILP) and the GA will be evaluated in this part. To this end, the model and the proposed algorithm will be applied to a real case problem, and the results will be compared to the outcomes of the other methods.

MATLAB® software v2016a was used for programming the proposed algorithms. Also, a laptop with an Intel® Core i7 6500U 2.5 GHz processor was used for testing the proposed algorithm.
4.1 Case study

The case study includes three building projects which are assigned to a GC. Tables 1, 2, 3, 4 and 5 summarize the information of these projects. It must be mentioned that $d_{ijk}$ and $d_{ijk}$ were considered the same in the case study problem.

4.2 The performance evaluation of the proposed model

The first evaluation performed in the current study is devoted to the proposed model. Figure 4 demonstrates the schedule results of the case study using the proposed model. In order to solve the model, the CPLEX solver of GAMS was employed.

In order to validate the obtained results, the model of Biruk et al. (2017) (which was proposed when only one project must be done) was also applied to the real case problem and its results were compared to the results of the proposed model. The results are summarized in Table 6. As is observed, the proposed model reduced the expenditure of the GC by around 14 percent.

Therefore, it can be concluded that the application of the model of Biruk et al. (2017) cannot result in an optimum solution in multiple projects. These outcomes were predictable because Biruk et al. did not consider the required assumptions when numerous projects need to be done by the GC simultaneously. As a result, the proposed model suits better when numerous projects need to be done at the same time.
| Work package | Project | Employed SC | Full-timely | Subcontracting |
|--------------|---------|-------------|-------------|---------------|
|              |         |             | Week 23     | Week 24       |
|              |         |             |             |               |
|              |         |             | Week 25     | Week 26       |
|              |         |             |             |               |
|              |         |             | Week 27     | Week 28       |
|              |         |             |             |               |
|              |         |             | Week 29     | Week 30       |
|              |         |             |             |               |
|              |         |             | Week 31     | Week 32       |
|              |         |             |             |               |
|              |         |             | Week 33     | Week 34       |
|              |         |             |             |               |
|              |         |             | Week 35     |               |

T1=217, T2=186, T3=219
Total cost=20401

Fig. 4 continued
4.3 Performance of the proposed GA for subcontractor selection problem

4.3.1 Tuning the parameters

In order to evaluate the proposed GA for solving the subcontractor selection and allocation problem, first, the parameters of the proposed GA must be adjusted. For this aim, the Taguchi experiment design method (Montgomery 2005) is used. In this regard, three levels for GA parameters (4 factors) are considered (see Table 7) that should be tested according to values which are illustrated in Table 8.

The response variable (RV) value (Eq. 31) is employed to compare the results of the experiments.

\[
RV = \frac{\text{Total cost}_{GA} - \text{OPT}_C}{\text{OPT}_C}
\]  

(31)

where Total cost_{GA} represents the case study total cost which is obtained through GA, and OPT_C is the optimal total cost of the case study which is calculated through the use of CPLEX solver of GAMS. The average results for each factor level have been reported in Fig. 5. As is observed, the optimum level of the factors A, B, C, and D are A(1), B(3), C(1), and D(2).

4.3.2 Performance of the proposed heuristic to produce initial population

In order to test the performance of the proposed heuristic for the production of the initial population, we compare its results with the random generation method. For this purpose, each of the procedures was applied 10 times on the case study. The experiments show that the results are improved by 10 percent using the proposed heuristic, while it increases CPU time 3 percent.

4.3.3 Performance of the proposed fitness function

In order to compare the proposed fitness function with a fitness function that works based on removing infeasible solutions, each of them was applied 10 times in the case study. The results in Table 9 show that the proposed fitness function improves the desirability of the solutions by 19%, while it increases CPU time slightly.

4.3.4 Performance of the proposed GA

Average deviation and run time as two performance measurement indices are utilized to evaluate the performance of proposed GA for solving the subcontractor selection and allocation in multiple building projects. Similar to previous performance evaluation, in order to test the performance of the proposed GA, it was run 10 times to solve the case study problem. The results are presented in Table 10. Optimum values for the case study which are computed by the CPLEX solver of GAMS are considered as the basis for the calculation of the average deviations.

It can be observed that the proposed algorithm can reduce computing time by 8% in comparison with the exact method and at the same time it is capable of reaching desirable solutions. In fact, although the exact method can probe the search space completely, the searching time will

---

Table 7 Definition of factor levels for parameters of GA

| Factor   | Symbol | Number of levels | Level definition |
|----------|--------|-----------------|-----------------|
| POP_size | A      | 3               | A1 : 50         |
|          |        |                 | A2 : 75         |
|          |        |                 | A3 : 100        |
| ν        | B      | 3               | B1:1            |
|          |        |                 | B2:2            |
|          |        |                 | B3:3            |
| Pc       | C      | 3               | C1 : 0.20       |
|          |        |                 | C2 : 0.50       |
|          |        |                 | C3 : 0.80       |
| Pm       | D      | 3               | D1 : 0.15       |
|          |        |                 | D2 : 0.20       |
|          |        |                 | D3 : 0.25       |

Table 8 The orthogonal arrays L9

| Experiment number | A | B | C | D |
|-------------------|---|---|---|---|
| 1                 | 1 | 1 | 1 | 1 |
| 2                 | 1 | 2 | 2 | 2 |
| 3                 | 1 | 3 | 3 | 3 |
| 4                 | 2 | 1 | 2 | 3 |
| 5                 | 2 | 2 | 3 | 1 |
| 6                 | 2 | 3 | 1 | 2 |
| 7                 | 3 | 1 | 3 | 2 |
| 8                 | 3 | 2 | 1 | 3 |
| 9                 | 3 | 3 | 2 | 1 |

---
increase significantly. So, adopting this approach seems to be unreasonable in large-scale problems because the problem is an NP-hard problem. Plus, as the inputs increase, the outputs increase exponentially. Clearly,

![Graph A](image1)

![Graph B](image2)

![Graph C](image3)

![Graph D](image4)

**Fig. 5** The average VR for GA

**Table 9** Performances of the proposed fitness function and the fitness function that works based on removing infeasible solutions

| Performance index       | Proposed fitness function | Fitness function that works based on removing infeasible solutions |
|-------------------------|---------------------------|------------------------------------------------------------------|
| Total cost              | 980.25 $                  | 1205 $                                                           |
| Calculation time        | 13.25 min                 | 12 min                                                           |

**Table 10** Performance of the proposed GA

| Performance measurement index | GA    | CPLEX solver of GAMS |
|-------------------------------|-------|----------------------|
| Average deviation             | 980.25 $ | 0                    |
| Run time                      | 13.25 min | 170 min              |
metaheuristic approaches like GA can result in desirable results in a reasonable amount of time.

5 Managerial implications

The current research study presents a practical tool for GC to the management of multiple projects in four ways:

- Cost optimization GCs as economical enterprises are always looking to reduce costs. The current study provides a practical tool for this purpose by considering the real conditions of building projects.
- Time Management In order to maintain the credibility of the main contractor with the employer, the main contractor wants its projects to be completed within the deadlines specified in the contract. This paper presents a software for scheduling multiple projects such that the mentioned goal is achieved.
- Portfolio management GCs generally conduct some projects simultaneously. They classify their projects into several categories such that the projects with the same resources are located in the same categories. Because the organization’s resources are limited, these projects should be prioritized for resource allocation. This is performed based on GCs 's goal that is usually cost minimization. This study is provided a practical for this purpose.
- Subcontractor selection This paper creates a software for managing subcontractors. In this regard, it identifies the right subcontractors and assigns the available work packages of a project to them.

6 Conclusion

This research study proposes a real model for subcontractor selection and allocation in multiple building projects. In this regard, the required assumptions are discussed and then a mixed-integer linear programming (MILP) model is suggested. In this model, different locations, times, and effort which is required for the movement of employed skilled workers and subcontractors from one project to another are considered. Then, the model will be solved by means of an exact method and a metaheuristic method. The contributions of the proposed algorithm are threefold: (1) a heuristic method is presented for the generation of the initial population; (2) an infeasible tackling procedure is proposed in order to change infeasible solutions to feasible ones; and (3) a simple fitness function is suggested to handle feasible and infeasible solutions. The results show that:

1. When numerous projects need to be done at the same time, single project planning criteria will not result in an optimum solution.
2. The proposed MILP model also has better performance in multiple project environments compared to the existing models.
3. The proposed heuristic method has better performance (approximately 10%) in reaching better solutions.
4. The proposed fitness function could improve the desirability of the produced feasible solutions by 19% (at, respectively, the same as the approach of removing infeasible solution).
5. The proposed GA in stopping criteria of 5000 results in lower computing time (92% in comparison with the exact method). Besides, it can reach the desirable solutions.

Author contributions MRA, VSV, and MHS contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

Funding The authors have no relevant financial disclosures.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Afshar MR, Shahhosseini V, Sebt MH (2020a) An interval type-2 fuzzy MCDM model for work package subcontractor prequalification. Soft Comput 25:1–14
Afshar MR, Shahhosseini V, Sebt MH (2020) Optimal sub-contractor selection and allocation in a multiple construction project: project portfolio planning in practice. J Oper Res Soc 1–14.
Afshar MR, Alipouri Y, Sebt MH, Chan WT (2017) A type-2 fuzzy set model for contractor prequalification. Autom Constr 84:356–366
Agrawal S, Sarkar S, Srivastava G, Maddikunta PKR, Gadekallu TR (2021) Genetically optimized prediction of remaining useful life. Sustain Comput Inform Syst 31:100565
Beşikçi U, Bilge Ü, Ulusoy G (2015) Multi-mode resource constrained multi-project scheduling and resource portfolio problem. Eur J Oper Res 240(1):22–31
Biruk S, Jaśkowski P, Czarnigowska A (2017) Minimizing project cost by integrating subcontractor selection decisions with scheduling. In: IOP conference series: materials science and engineering. IOP Publishing, p. 072007
El-Abbasy MS, Elazouni A, Zayed T (2017) Generic scheduling optimization model for multiple construction projects. J Comput Civ Eng 31(4):04017003
Fachurrazi HS, Munirwansyah H (2017) The subcontractor selection practice using ANN-multilayer. Int J Technol 8(4):761–772
Józefowska J, Mika M, Różycki R, Waliłgórą G, Węglarz J (2001) Simulated annealing for multi-mode resource-constrained project scheduling. Ann Oper Res 102(1–4):137–155
Khan MU, Javed AR, Ihsan M, Tariq U (2020) A novel category detection of social media reviews in the restaurant industry. Multimed Syst 1–14

Lee K, El-Sharkawi M (2008) Modern heuristic optimization techniques: theory and applications to power systems. IEEE press series on power engineering. Wiley-IEEE Press, New York

Montgomery DC (2005) Design and analysis of experiments, 6th edn. JohnWiley & Sons, Hoboken

Mokhtari H, Abadi INK (2013) Scheduling with an outsourcing option on both manufacturer and subcontractors. Comput Oper Res 40(5):1234–1242

Monghasemi S, Nikoo MR, Fasaee MAK, Adamowski J (2015) A novel multi criteria decision making model for optimizing time-cost-quality trade-off problems in construction projects. Expert Syst Appl 42(6):3089–3104

Mungle S, Benyoucef L, Son YJ, Tiwari MK (2013) A fuzzy clustering-based genetic algorithm approach for time-cost-quality trade-off problems: a case study of highway construction project. Eng Appl Artif Intell 26(8):1953–1966

Pells D (1993) The AMA handbook of project management AMA-COM, New York

Polat G (2016) Subcontractor selection using the integration of the AHP and PROMETHEE methods. J Civ Eng Manag 22(8):1042–1054

Polat G, Kaplan B, Bingol BN (2015) Subcontractor selection using genetic algorithm. Procedia Eng 123:432–440

Pollack-Johnson B, Liberatore MJ (2006) Incorporating quality considerations into project time/cost tradeoff analysis and decision making. IEEE Trans Eng Manage 53(4):534–542

Rehman SU, Javed AR, Khan MU, Nazar Awan M, Farukh A, Hussien A (2020) PersonalisedComfort: a personalised thermal comfort model to predict thermal sensation votes for smart building residents. Enterprise Inf Syst 1–23

Sonmez R, Gürel M (2016) Hybrid optimization method for large-scale multimode resource-constrained project scheduling problem. J Manag Eng 32(6):04016020

Ulubeyli S, Kazaz A (2016) Fuzzy multi-criteria decision making model for subcontractor selection in international construction projects. Technol Econ Dev Econ 22(2):210–234

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.