A New Optimization Model for Reliable Team Formation Problem Considering Experts' Collaboration Network

Mohammad Fathian, Mohamad Saei-Shahi, and Ahmad Makui

***Abstract*—The objective of this study is to propose a new optimization model for the formation of a reliable team of experts, who have a certain number of skills and best collaboration with each other. The proposed mathematical model maximized team reliability by considering the probability of unreliable experts that may leave the team with the probability (1-Q) and proposed a backup for each unreliable member. In this paper, a new model was developed for simultaneously forming a team with three key factors: 1) expert's skills; 2) expert's collaboration network; and 3) expert's reliability. This model was evaluated by two numerical studies on both artificial and real-life datasets with small and big data. The optimum combination of team members with regard to skills, collaboration network, and reliability will help managers in performing their projects or operations. Therefore, the new optimization model, called the reliable team formation problem, proposed team members in two sets consisting of main and backup members.**

***Index Terms*—Collaboration network, operations research, optimization, reliability, team formation.**

I. Introduction

HUMANS have always tended to meet their needs and aspirations through group activities. The formation of groups and teams in the past has been subjected to geography and local constraints. Indeed, people whose residences were within the same proximity formed teams. Obviously, in such circumstances, the provision of all required capabilities and resources was not always possible, and people were given tasks for which they lacked the necessary skills.

Today, with regard to the trend of societies to specialization, the emergence of high-speed vehicles and communication via the internet, the new issue in the formation of a team now involves the bringing together of people who can cover all needed skills from different geographical locations. Several researchers have addressed this issue [1], [2]. However, in order to form the most effective team, the level of collaborations among team members would be a challenging issue. To illustrate this problem, let us consider a research team that includes the best experts from various spheres of knowledge, who work on an emerging technology. Here, in addition to all the necessary specialties related to that technology, the success of the team depends on proper collaborations among team members. Thus, the mentioned problem could be briefly stated as follows: determining a subset of available experts to cover required skills in a way that they can best collaborate with each other to promote projects and activities.

Lappas *et al.* [3] focused on the importance of collaborations between experts by taking the cost of the collaborations into account in their team formation model, in order to measure the level of collaborations. In their model, it was assumed that all experts would be available from the beginning to the end of the project. However, in reality, some members may leave after team formation, which leads to a crucial problem in the continuation and execution of the project. The problem could be extremely complex because the new alternative should have a pleasant history of collaboration with other team members. In order to cope with this significant problem, this paper has developed a model for the formation of a reliable team of experts, considering their collaboration network (that is, a network includes the level of collaboration between each pair of experts according to their work history). Indeed, the proposed model is aimed at finding a reliable team of experts who excellently collaborate with each other, and covers a certain number of skills.

The reliability of working teams is one of the serious issues in team formation, and it entails covering the required activities as well as good cooperation among team members. Therefore, the effort of project managers and team leaders to form a team which has appropriate sustainability must include the coverage of required activities and the promotion of good collaboration among the members. The occurrence of disruption is one reason for the nonsustainability of teams. One of the most common disruptions in teams is caused by the quitting of team members during projects. The effect of quitting a team becomes greatly felt when the team members have had a good cooperation and if the new member finds it difficult to properly collaborate with other members; thus, leading to a failure of the project. Therefore, in some cases, this issue had to be compensated for by reforming the entire team. Therefore, establishing a team which has more reliability and whose team members have a maximum collaboration is of great necessity.

Nonattendance by experts to the end of the activity may be caused by various reasons, for example, illness, family

Manuscript received September 22, 2016; revised February 11, 2017; accepted May 30, 2017. Review of this manuscript as arranged by Department Editor P. ED Love. *(Corresponding Author: Mohammad Fathian.)*

The authors are with the School of Industrial Engineering, Iran University of Science and Technology, Tehran17697, Iran (e-mail: fathian@iust.ac.ir; saeishahi@ind.iust.ac.ir; amakui@iust.ac.ir).

Digital Object Identifier 10.1109/TEM.2017.2715825

emergencies, taking a new job, etc. Some of these may be unpredictable and occur during projects; hence, it is difficult to measure the risk of these events before forming a team. Therefore, this study focused on measurable risk. One way of measuring this risk probability is by the self-assessment of experts.

Another method emphasized by this study is the use of a historical data of expert activities. This measure can be defined by the ratio of the number of successful projects of expert (when experts did not leave the team) to the number of all expert's projects. In real problems, each person may have a different probability, but in the first step to develop a linear mathematical model, this probability was assumed equal for all experts.

One of the methods of establishing a reliable team is by considering a backup for each member of the team. This backup serves as a replacement incase a member decides to quit the team suddenly.

The selection of a backup when available experts are limited is a complex problem in itself.

Although numerous studies have considered the issue of team cooperation in addition to covering skills in teams, none talked about the formation of a reliable team and selection of a backup. In simple terms, a good team may be formed by considering expert skills and their social networks, as well as suggest backups after that, but one important question still remains; when the backup person replaces the expert who left the team, is the new network the most appropriate one? Obviously, the roles of people in a sustainable and efficient team (and even near-optimal) are different. So, there is a deep gap in researches and this addresses the issue of the possibility of selected individuals' presence to the end of the team activity.

In this paper, to cover the problem, a mathematical model was presented. The model focused on finding a reliable team of experts, who in addition to providing a certain number of skills, formed the best collaboration network.

The remaining parts are as follows. The next section reviews the literature. Section III describes the team formation model based on the aforementioned requirements. Section IV provides the numerical experiments and the respective results, while Section V presents the concluding remarks.

II. Literature Review

The problems associated with the formation of a team refer to the selection of a subset of available skillful individuals, so that they can cover a predefined set of skills. It is not far from mind that a significant number of studies in this area talked about the field of operations research [1], [4], [5]. Formulating the team formation problem as an integer linear programming and then focusing on finding optimal matching between people and performance demanding requirements was the first approach developed in the team formation problem from the perspective of the optimization field. Since this problem can be reduced to the covering problem, it cannot be solved within a short space of time. Therefore, it is necessary to use common types of search algorithms, such as the heuristic and meta-heuristic algorithms. Thus, in literature, this problem has been repeatedly solved by techniques such as simulated annealing [6], the Branch and Cut algorithm [5], as well as the Genetic algorithm [7]. Some studies have focused on skills and their requirements in team work [8], [9].

In real life, people who work as a team cannot be considered as machines, which are only connected with each other to perform tasks. They are human beings with human characteristics. Thus, to form a proper team, it is essential to consider the nature of individuals; so some researchers considered the importance of effective cooperation among members in a team. Chen and Lin [1] used the Myers–Briggs test for personality measurement and internal communications assessment among team members. One year later, in another study, Fitzpatrick and Askin [10] used the KCI[1] index to measure the temperament and nature of team members.

The subject of the communication structure of people in the team was mentioned for the first time by Gaston *et al.* [11]. The authors of this paper were able to demonstrate that adjusting communication structure between individuals, because of its effect on team formation within an organization, greatly affected organizational performance.

The problem of team formation was first brought up by Lappas *et al.* [3], by emphasizing the importance of considering the social networks of experts and relation costs among them. They offered two heuristic methods based on the "Steiner tree," called "Cover-Steiner" and "Enhanced-Steiner," to solve the team formation problem, when any activity needs only one person for any skill.

Research studies conducted in 2010 and 2011, respectively, defined an objective function to allocate activities to experts when forming teams [6] and solved the problem of finding the best k team of experts [12], while Datta *et al.* [13] considered capacity constraints in their model in a way to ensure that no expert was allocated an activity which was more than their capacity. In fact, they regarded the capacities of experts as limitations instead of covering skills, and tried to minimize the cost of social communication. Another study related to team formation was conducted by Sorkhi *et al.* [9]. They carried out team formation by giving the costs associated with the different levels of the experts. Cheng *et al.* [14] considered the number of experts required by using the generalized Enhanced–Steiner algorithm, and also presented a grouping-based method, by condensing the expert information into a compact representation group graph based on the required skills.

Farasat and Nikolaev [15] developed a mathematical framework that relies on modern social network analysis theories for solving problems associated with nurse team formation and nurse scheduling (shift assignment), thereby accounting for signed social connections.

So far, the literature category about team formation has not been provided. Therefore, this paper proposed a general category for the literature; two main approaches can be obtained for the team formation literature. The first is the skill-driven approach, while the second is the relationship-driven approach.

[1] Kolbe Conative Index.

TABLE I
CATEGORIES OF TEAM FORMATION USING THE COLLABORATION NETWORK OF EXPERT

Issue	Paper
Communication Cost	[2], [3], [9], [12]–[14], [16]–[28]
Number of Skilled Experts	[14], [29]
Capacity of Experts	[13]
Balance of Workload	[2], [6]
Team Leader	[5], [12], [22], [26], [30], [31]
Skillfulness Level of Experts	[9], [16], [32]
Skills and Connectivity Trade-Off	[33]–[36]
Multiple Tasks	[2], [37]
Personal Cost	[17], [38]
Locational	[22], [24], [27], [39], [40]
Organization Network	[11], [41], [42], [43], [44], [45], [46], [47]
Clustering	[16], [48], [49]
Time Constraint	[50], [51]
Diversity	[11], [18], [46], [52], [53]
Dynamics	[52], [54], [55]

A. Skills-Based Approaches

Skills-based approach refers to methods and algorithms which are based on satisfying the required skills. In this approach, however, other issues have been raised from the issue of skills. However, what algorithms were relied on and emphasized is the need for providing the necessary skills to carry out activities.

The research studies conducted according to this approach monitored the development of a mathematical model in the field of operations research and made use of meta-heuristic algorithms, such as Simulated Annealing [4] and Genetic Algorithms [7].

B. Relationship-Based Approaches

Although the methods in this section, which are often seen in recent articles, were used to find people which covered the required skills, but first of all they emphasized on finding a sub graph with a low cost of relationship between the social network of experts [3]. A category of issues regarding the relationship-driven approaches in team formation is shown in Table I.

Table I presents 15 important issues mentioned in the team formation literature with relationship-driven approaches. Communication cost, number of skilled experts, capacity of experts, balance of workload, team leader, skillfulness level of experts, skills and connectivity trade-off, multiple tasks, personal cost, locational, organization network, clustering, time constraint, diversity and dynamics are the main issues that researchers have considered.

As shown in Table I, "Communication Cost" and "Organization Network" are more interesting to researchers.

According to Table I, it is clear that the combination of issues has been addressed, but a significant issue that has not been discussed in the team formation literature is the formation of a reliable team.

III. PROPOSED MATHEMATICAL MODEL

This section describes the proposed mathematical model used in solving team formation problem where reliability emerges as a crucial issue. This model is called a reliable team formation problem (RTFP).

Given a set of Experts I as candidates for forming a team, we have a set of positions of team members, called P, which shows the total number of team members. Moreover, there is a set of required skills, called S, for each of the candidates. Each expert would be represented as a node and the relationship between them would be illustrated by an edge. In RTFP, there are two sets of decision variables: 1) the assignment variables; and 2) the relationship variables. The assignment variables are denoted by x_i, which equals 1 if expert i is assigned to the team, while the relationship variables are denoted by y_{pi}, which equals 1 if expert i is assigned to the position p.

Two groups of experts are assumed: 1) experts who are reliable and will remain in the team to the end of the project; and 2) experts who are unreliable and may leave the team with the same probability $(1 - Q)$, independently. The first set is called R (that is, Reliable) and the second set is called U (that is, Unreliable). Therefore, there is the need to define two column matrix sets A_i^R and A_i^U, which shows that each expert is reliable or unreliable. In set A_i^R, cell i equals 1 if expert i is reliable and in set A_i^U, cell i equals 1 if expert i is unreliable.

For Each position in the model, two types of members are considered. The first type refers to the main members of the team, denoted by M, and the second type refers to the backup members, indicated by B. Thereafter, two variables are defined for these two types: y_{pi}^M, which equals 1, if the expert i is assigned to the position p as a main member, and y_{pi}^B, which equals 1, if the expert i is assigned to the position p as a backup member.

Here, it should be noted that if a reliable person is assigned to the team as a main member in a position, then there is no need to assign another one for backup in the respective position, and if an unreliable one is assigned to the team as a main member in a position, there is need to assign only another unreliable one as a backup in that position. Thus, in the main positions, both reliable and unreliable persons are assigned. However, in the backup positions, only unreliable persons are assigned.

According to the type of expert (reliable or unreliable) and the type of team member (main or backup), there are six types of relationships.

Six binary variables could be considered for the relationships. The first variable is denoted by rm_{ij}^{RR}, which equals 1 if there is a relationship between two main reliable team members. The second variable is denoted by rm_{ij}^{RU}, which equals 1 if there is a relationship between one main reliable team member and another main unreliable team member. The third variable is denoted by rm_{ij}^{UU}, which equals 1 if there is a relationship between two main unreliable team members. The fourth variable is denoted by rb_{ij}^{UU}, which equals 1 if there is a relationship between two backup unreliable team members. The fifth variable is denoted by rmb_{ij}^{RU}, which equals 1 if there is a relationship between one main reliable team member and another backup unreliable team member, and the last variable is denoted by rmb_{ij}^{UU}, which equals 1 if there is a relationship between one main unreliable team member and another backup unreliable team member.

In the present model, the parameter SA_{si} denotes the assignment of skill s for expert i, and PS_{ps} that denotes the required skills for each position. The weight of relationship between each pair of experts is defined by $W_{ij}\ \forall i.j \in I$. Finally, K_s is considered as the minimum number of each skill required for a team.

A. Objective Function

The objective function (1) maximizes the total weight of the relationships between the team members. The probability of remaining the Reliable experts in the team equals 1, and in the case of unreliable experts the probability equals Q

$$\begin{aligned}&(\text{RTFP})\\ &\text{Max} \sum_{i\in I}\sum_{j\in J} W_{ij} rm_{ij}^{\text{RR}} + \sum_{i\in I}\sum_{j\in J} Q\, W_{ij}\, rm_{ij}^{\text{RU}}\\ &\qquad + \sum_{i\in I}\sum_{j\in J} Q^2 W_{ij} rm_{ij}^{\text{UU}}\\ &\qquad + \sum_{i\in I}\sum_{j\in J} Q^2 (1 - Q)^2 W_{ij} rb_{ij}^{\text{UU}}\\ &\qquad + \sum_{i\in I}\sum_{j\in J} Q\,(1 - Q)\, W_{ij} rmb_{ij}^{\text{RU}}\\ &\qquad + \sum_{i\in I}\sum_{j\in J} Q^2\,(1 - Q)\, W_{ij} rmb_{ij}^{\text{UU}}. \quad (1)\end{aligned}$$

As earlier stated, there are six types of relationships in the present model. With two experts selected as members of the team simultaneously, the probability of their relationship in the calculations will depend on the type of these members (that is, reliable or unreliable and main or backup). If both are main and reliable, the probability equals 1. If these two members are main and only one of them is reliable, then the probability equals $1 \times Q$. If both of them are main and unreliable, the probability is equal to Q^2. If both of them are backup and unreliable, this implies that the respective main members may leave the team; therefore, the probability would be calculated as $Q^2(1 - Q)^2$. If one of the members is the main and is reliable, while the other is the backup and is unreliable, because of the absence of the main member in the position of the backup member, the probability equals $1 \times Q \times (1 - Q)$. If one of the members is the main and is unreliable while the other is the backup and is unreliable, because of the absence of the main member in the position of the backup member, the probability equals $Q \times Q \times (1 - Q)$. Table II shows a clear perception of these relationships and their probabilities.

B. Constraints

In the RTFP model, 15 constraints are defined which can be described in three categories: assignment constraints, Relationship constraints, and Skill constraints.

TABLE II
RELATIONSHIPS PROBABILITY

Member 1	Member 2	Related var.	Probability
Main-Reliable	Main-Reliable	rm_{ij}^{RR}	1
Main-Reliable	Main-Unreliable	rm_{ij}^{RU}	$1{*}Q$
Main-Unreliable	Main-Unreliable	rm_{ij}^{UU}	Q^2
Backup-Unreliable	Backup-Unreliable	rb_{ij}^{UU}	$Q^2(1 - Q)^2$
Main-Reliable	Backup-Unreliable	rmb_{ij}^{RU}	$1{*}Q{*}(1 - Q)$
Main-Unreliable	Backup-Unreliable	rmb_{ij}^{UU}	$Q{*}Q{*}(1 - Q)$

1) Assignment Constraints:

$$\sum_{i\in I} y_{pi}^{M} = 1 \qquad \forall p \in P \qquad (2)$$

$$\sum_{i\in I} y_{pi}^{B} \le 1 \qquad \forall p \in P \qquad (3)$$

$$x_i = \sum_{p\in P} y_{pi}^{M} + \sum_{p\in P} y_{pi}^{B} \qquad \forall i \in I \qquad (4)$$

$$y_{pi}^{B} \le A_i^{U} \qquad \forall i \in I\ \&\ s \in S \qquad (5)$$

$$b_p = 1 - \sum_{i\in I} A_i^{R} y_{pi}^{M} \qquad \forall p \in P \qquad (6)$$

$$A_i^{U} y_{pi}^{B} \le b_p \qquad \forall i \in I\ \&\ p \in P. \qquad (7)$$

In the RTFP model, Constraint (2) states that one expert must be assigned as main for each position. Constraint (3) states that at least one expert should be assigned as backup for each position. Constraint (4) defines each candidate with (x_i) assigned as a team member. According to the assumptions of the model, all the members in backups must be unreliable; therefore, constraint (5) prevents the assignment of a reliable expert as a backup team member. It is assumed that if a reliable team member is assigned in one position, then the need for a backup for him is unnecessary, therefore constraint (6) prevents the assigning of experts in backup position if another reliable one is assigned as the main team member for that position. Our priority is to assign only reliable members as the main member; therefore, constraint (7) is defined to force the model to assign only an unreliable expert as backup.

2) Relationship Constraints:

$$2{*}rm_{ij}^{\text{RR}} \le A_i^{R} \sum_{p\in P} y_{pi}^{M} + A_j^{\text{R}} \sum_{p\in P} y_{pj}^{M} \quad \forall i,j \in I\ \&\ i<j \qquad (8)$$

$$2{*}rm_{ij}^{\text{RU}} \le A_i^{\text{R}} \sum_{p\in P} y_{pi}^{M} + A_j^{U} \sum_{p\in P} y_{pj}^{M} \quad \forall i,j \in I\ \&\ i<j \qquad (9)$$

$$2{*}rm_{ij}^{\text{UU}} \le A_i^{U} \sum_{p\in P} y_{pi}^{M} + A_j^{U} \sum_{p\in P} y_{pj}^{M} \quad \forall i,j \in I\ \&\ i<j \qquad (10)$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

$$2{*}rb_{ij}^{\mathrm{UU}} \leq A_i^U \sum_{p\in P} y_{pi}^B + A_j^U \sum_{p\in P} y_{pj}^B \qquad \forall i,j \in I \;\&\; i<j \tag{11}$$

$$2{*}rmb_{ij}^{\mathrm{RU}} \leq A_i^{\mathrm{R}} \sum_{p\in P} y_{pi}^M + A_j^U \sum_{p\in P} y_{pj}^B \quad \forall i,j \in I \;\&\; i<j \tag{12}$$

$$2{*}rmb_{ij}^{\mathrm{UU}} \leq A_i^U \sum_{p\in P} y_{pi}^M + A_j^U \sum_{p\in P} y_{pj}^B \quad \forall i,j \in I \;\&\; i<j \tag{13}$$

$$2 - rmb_{ij}^{\mathrm{UU}} \leq A_i^U y_{pi}^M + A_j^U y_{pj}^B \qquad \forall i,j \in I \;\&\; i < j \;\&\; p \in P. \tag{14}$$

As shown in Table II, there are six types of relationships between two members in the RTFP model. Therefore, constraints (8)–(14) are defined for these relationships. As shown in this model, there is an another constraint to defining the relationship between one main-unreliable and another backup-unreliable team member. According to the assumptions of this model, it is unnecessary to consider their relationships in the same position, because when the main member leaves the team, the backup member works in the same position; hence, the main and his backup do not exist simultaneously in a team. Thus, in considering this issue, constraint (14) is defined.

3) Skill Constraints:

$$K_s \leq \sum_{i\in I} x_i SA_{si} \quad \forall s \in S \tag{15}$$

$$PS_{ps} \leq \sum_{i\in I} SA_{si} y_{pi}^M \quad \forall s \in S \;\&\; p \in P \tag{16}$$

$$PS_{ps} \leq \sum_{i\in I} SA_{si} (y_{pi}^B + A_i^{\mathrm{R}} y_{pi}^M) \quad \forall s \in S \;\&\; p \in P. \tag{17}$$

In team formation, the team builder may define the minimum number of members' requirement for each skill; therefore, Constraint (15) is defined to force the model in considering this limitation. Another issue which must be defined by the team builder before team formation is the required skills for each position by PS_{ps}. Thus, Constraint (16) considers this parameter for team formation and assigns team members with specific skills, defined by PS_{ps} in the main team members, while constraint (17) considers this limitation for backup team members.

As described in Section III, all ten variables in this model are binary.x_i, y_{pi}^M, y_{pi}^B, b_p, rm_{ij}^{RR}, rm_{ij}^{RU}, rm_{ij}^{UU}, rb_{ij}^{UU}, rmb_{ij}^{RU}, $rmb_{ij}^{\mathrm{UU}} \in \{0,1\}\; \forall i,j \in I \;\&\; p \in P$.

IV. Experimental Results

In order to investigate the effectiveness and evaluate the performance of the proposed model, two methods can be used. The first method is to evaluate the model by using prepared or simulated datasets and evaluating their results.

The second method is to evaluate the performance of the model in a real environment. By so doing, several project teams can be formed in a particular organization for a short time by using the developed model and then we can compare the teams working performance with similar project teams in that organization.

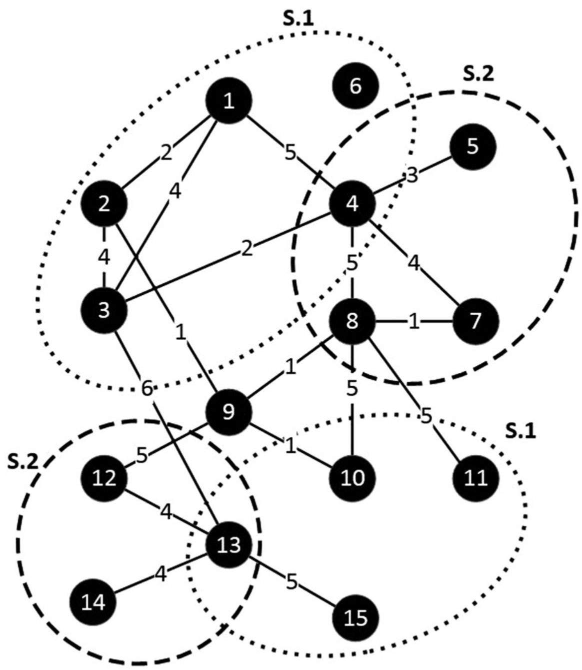

Fig. 1. Expert network with their skills.

In order to make sense of how the model works and evaluate the model, the first method is used. Therefore, this section provided two numerical experiments. The first experiment employed a small artificial network of specialists with limited skills. The second experiment employed a large set of experts using the real-life DBLP [56] dataset. Thereafter, the model was run on both datasets, using the GAMS[2] software [57] on a server with 30 core, 1.8 GHz, and 32-GB RAM, and the results of running the model for each problem is presented in the respective sections.

A. Artificial Network

1) Dataset: A set of 15 experts with two skills was considered, with each person having skill 1, skill 2, and both skills 1 and 2 or no skill.

Some of these experts have worked together in the past. This collaboration is shown in Fig. 1 by edges and nodes. The nodes represent the experts, while the edges represent the previous collaboration between them. The numbers shown on the edges represent the weight of collaboration between them, such that the higher number indicates further cooperation between them. Each expert's skills are shown in the figure. S.1 and S.2, respectively, represent the people with skill 1 and the people with skill 2, and show each individual skill. Some experts were reliable while others were unreliable. Therefore, Matrix A_i^{R} denotes the reliable experts while matrix A_i^U denotes the unreliable experts. Matrix SA_{si} denotes each individual skill. In this problem,

[2]General Algebraic Modeling System.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

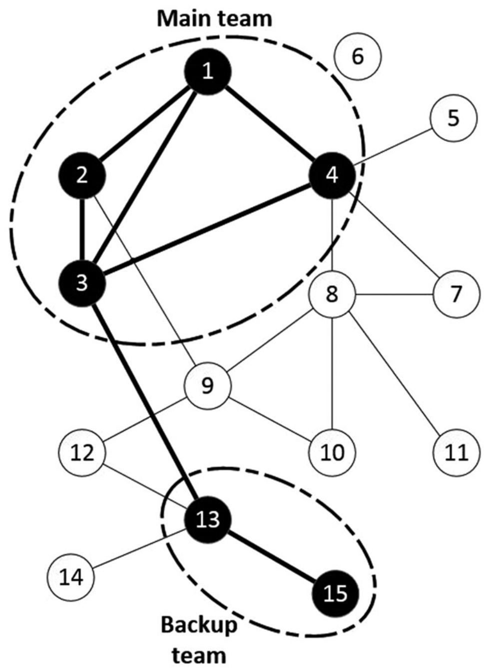

Fig. 2. Best reliable team.

0.6 defines the probability of the remaining unreliable experts

$$SA_{si} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$
$$A_i^{\mathrm{R}} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
$$A_i^{U} = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Here, the aim was to form a reliable team with 4 members from 15 experts. Four positions were defined for each member with each member satisfying specific skills. Matrix PS_{ps} shows the required skills for the position of each member.

$$PS_{ps} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

2) Results: After solving the proposed RTFP model, the most reliable team was obtained. The main team, members according to their position, was $y_p^M \in [x_2 \; x_4 \; x_3 \; x_1]$. Two members of the team were reliable (that is, x_1, x_2) and the others were unreliable (that is, x_3, x_4). According to the assumption of the model, for each unreliable member a backup member should be assigned. Therefore, the model proposed two backups as $y_p^B \in [- \; x_{13} \; x_{15} \; -]$ according to the position x_3, x_4. Hence, all main and backup team members were $x_i \in [x_1 \; x_2 \; x_3 \; x_4 \; x_{13} \; x_{15}]$ and the objective function in this problem was equal to $Z = 11.627$.

Figs. 2 and 3 show the results of the model in the case of the artificial network. As shown in Fig. 3, expert 13 was proposed as the backup for expert 4 and expert 15 was considered as the backup for expert 3.

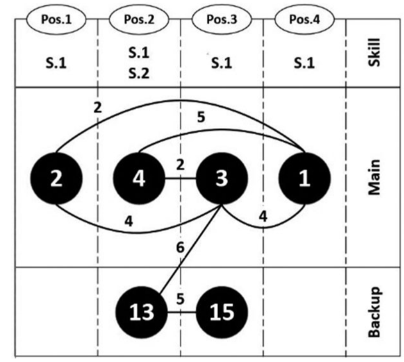

Fig. 3. Reliable team with their types and positions.

Fig. 3 presents the rearranged view of the final team and shows the position of each team member. The required skills for each position is as presented at the top of the figure. In the middle, the main members of the team are shown and at the bottom of the figure, the backup of each member is presented.

In this team, there existed seven important relationships which affected the result. These relationships are $rm_{1,2}^{\mathrm{RR}}$, $rm_{1,3}^{\mathrm{RU}}$, $rm_{1,4}^{\mathrm{RU}}$, $rm_{2,3}^{\mathrm{RU}}$, $rm_{3,4}^{\mathrm{UU}}$, $rb_{13,15}^{\mathrm{UU}}$, $rmb_{3,13}^{\mathrm{UU}}$.

This result showed that all assumptions in this paper were covered by the RTFP model. This assumption included: 1) the high priority of selection of existing reliable experts if any (x_1, x_2); 2) no backup assigned for reliable experts; 3) assigning of backups for unreliable experts from another unreliable one (x_{13}, x_{15} for x_3, x_4); and 4) assigning a backup with good cooperation with other main members ($w_{3,13} = 6$).

B. Real-Life Network

In order to evaluate the model using real data, the DBLP database was used as the dataset, which is publicly available through the DBLP portal.

1) Dataset: This data set has also been used by other researchers in team formation [14], [16], [3], [17]. The data used in this article are a part of the research data of Farhadi *et al.* [16]. The data were collected on April 12, 2006, and contained information of published papers at conferences in the areas of databases, data mining, artificial intelligence, and several other related issues. The papers were obtained by the following known conferences in relation to the aforementioned fields: the International Conference on Management of Data, the International Conference on Very Large Data Bases, the International Conference on Data Engineering Workshops, the International Conference on Database Theory, the International Conference on Extending Database Technology, the Symposium on Principles of Database Systems, the World Wide Web Conference, the Conference on Knowledge Discovery and Data Mining, the SIAM International Conference on Data Mining, the European Conference on Principles of Data Mining and Knowledge Discovery, the IEEE International Conference on Data Mining, the Symposium on Theoretical Aspects of Computer Science, the ACM Symposium on Theory of Computing, the IEEE Annual Symposium on Foundations of Computer Science, the ACM-SIAM

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

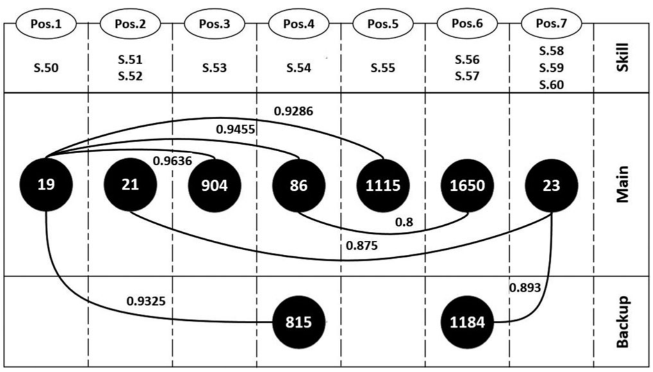

Fig. 4. Best obtained team with their types and positions.

Symposium on Discrete Algorithms, the Association for Uncertainty in Artificial Intelligence, the Conference on Learning Theory, the European Conference on Machine Learning. The skill set of each of the authors included terms and phrases which appeared at least in two titles of their articles in DBLP. Two authors were connected to each other regarding collaboration experience in at least two joint papers. The relation weight or the same weight on connection edges between two authors can be calculated by

$$W_{ij} = \frac{|p_i \cap p_j|}{|p_i \cup p_j|} \tag{18}$$

where p_i and p_j are the number of papers published by authors i and j, respectively. There are 2067 authors, 1000 distinct skills, and 4195 edges (collaboration links) in this dataset.

Here, the aim was to form a team with seven members from 2067 experts who covered 11 required skills (skills 50–60). In this problem, seven positions were defined for each member such that each member should satisfy specific skills. Matrix PS_{ps} shows the required skills for the position of each member.

Due to a lack of reliability data, experts were randomly assigned as reliable or unreliable and consider equal probability ($Q = 0.6$) for unreliable experts

$$PS_{ps} \in \begin{bmatrix} S_{50} & - & - \\ S_{51} & S_{52} & - \\ S_{53} & - & - \\ S_{54} & - & - \\ S_{55} & - & - \\ S_{56} & S_{57} & - \\ S_{58} & S_{59} & S_{60} \end{bmatrix} \qquad Q = 0.6.$$

2) Results: Fig. 4 presents the result of proposed RTFP model for this problem. The model formed a reliable team with seven members from 2067 experts. In this problem, seven experts were assigned according to their position as main team members $y_p^M \in [x_{19}\ x_{21}\ x_{904}\ x_{86}\ x_{1115}\ x_{1650}\ x_{23}]$ and two experts were assigned according to their position as backup team members $y_p^B \in [-\ -\ -\ x_{815}\ -\ x_{1184}\ -]$.

From the results and DBLP database, A_i^{R} are reliable while A_i^U are unreliable. Fig. 4 shows that unreliable expert 815 was assigned as a backup for unreliable expert 86 and unreliable expert 1184 was assigned as a backup for unreliable expert 1650. Every team member satisfied all the required skills. In this team, there were seven important relationships which affected the result. This relationship is given as $rm_{19,904}^{\mathrm{RR}}$, $rm_{19,1115}^{\mathrm{RR}}$, $rm_{21,23}^{\mathrm{RR}}$, $rm_{19,86}^{\mathrm{RU}}$, $rm_{19,86}^{\mathrm{RU}}$, $rm_{86,1650}^{\mathrm{UU}}$, $rmb_{19,815}^{\mathrm{UU}}$, $rmb_{23,1184}^{\mathrm{UU}}$. The goal function in this problem is equal to $\mathrm{Z} = 3.885$.

Based on the achieved solutions, it can be seen that all the assumptions made in this study were covered. Also, it can be seen that the first priority of selecting a member was from reliable experts and the model proposed the best backups for each unreliable one, having a good cooperation with others and having similar skills with their main members.

V. Conclusion

This paper discussed the problem of forming a reliable team of experts to cover a set of required skills, considered their collaboration network and introduced a mathematical model which forms a reliable team of experts. Unlike other proposed optimization models that often focus on people's skills, personality traits, abilities, and individual costs, this paper aimed to simultaneously consider the collaboration between experts and the reliability of team members. In doing so, the model attempted to obtain a proper backup for each unreliable team member.

To evaluate the model, an artificial numerical problem and a real problem with data available on the DBLP website was used. The experimental results confirmed that the proposed model can effectively form teams. One limitation of our model is that it considers equal probability for each unreliable expert, who may leave the team. However, in real life, this probability is unequal for each member. This limitation was retained as an interesting future work. One of the assumptions of this model is the availability of backups in the project duration. However, in the real world, a backup may not be available. In future studies, this assumption can be released and cost can be considered.

References

[1] S. Chen and L. Lin, "Modeling team member characteristics for the formation of a multifunctional team in concurrent engineering," *IEEE Trans Eng. Manage.*, vol. 51, no. 2, pp. 111–124, May 2004.

[2] A. Anagnostopoulos, L. Becchetti, C. Castillo, A. Gionis, and S. Leonardi, "Online team formation in social networks," in *Proc. ACM Int. Conf. World Wide Web*, 2012, pp. 839–848.

[3] T. Lappas, K. Liu, and E. Terzi, "Finding a team of experts in social networks," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Paris, France, 2009, pp. 467–476.

[4] A. Baykasoglu, T. Dereli, and S. Das, "Project team selection using fuzzy optimization approach," *Cybern. Syst.*, vol. 38, no. 2, pp. 155–185, 2007.

[5] H. Wi, S. Oh, J. Mun, and M. Jung, "A team formation model based on knowledge and collaboration," *Expert Syst. Appl.*, vol. 36, no. 5, pp. 9121–9134, 2009.

[6] A. Anagnostopoulos, L. Becchetti, C. Castillo, A. Gionis, and S. Leonardi, "Power in unity: Forming teams in large-scale community systems," in *Proc. 19th ACM Int. Conf. Inf. Knowl. Manage.*, 2010, pp. 599–608.

[7] A. Zakarian and A. Kusiak, "Forming teams: An analytical approach," *IIE Trans.*, vol. 31, pp. 85–97, 2004.

[8] G. Liu and S.-I. Yokoyama, "Proposal for a quantitative skill risk evaluation method using fault tree analysis," *IEEE Trans. Eng. Manage.*, vol. 62, no. 2, pp. 266–279, May 2015.

[9] M. Sorkhi, S. Hashemi, and A. Hamzeh, "An effective expert team formation in social networks based on skill grading," in *Proc. IEEE Int. Conf. Data Mining Workshops*, 2011, pp. 367–372.

[10] E. Fitzpatrick and R. Askin, "Forming effective worker teams with multi-functional skill requirements," *Comput. Ind. Eng.*, vol. 48, no. 3, pp. 593–608, 2005.
[11] M. E. Gaston, M. DesJardins, and J. Simmons, "Adapting network structures for efficient team formation," in *Proc. AAMAS-04 Workshop Learning Evolution Agent-Based Syst.*, Washington, DC, 2004.
[12] M. Kargar and A. An, "Discovering top-k teams of experts with/without a leader in social networks," in *Proc. ACM Int. Conf. Inf. Knowl. Manage.*, 2011, pp. 985–994.
[13] S. Datta, A. Majumder, and K. Naidu, "Capacitated team formation problem on social networks," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2012, pp. 1005–1013.
[14] L. Cheng, S. Man, and L. Shou, "On team formation with expertise query in collaborative," *Knowl. Inf. Syst.*, no. 42, pp. 441–463, Oct. 06, 2015.
[15] A. Farasat and A. G. Nikolaev, "Signed social structure optimization for shift assignment in the nurse scheduling problem," *Socio-Economic Planning Sci.*, vol. 56, pp. 3–13, 2016.
[16] F. Farhadi, E. Hoseini, S. Hashemi, and A. Hamzeh, "TeamFinder: A co-clustering based framework for finding an effective team of experts in social networks," in *Proc. 12th IEEE Int. Conf. Data Mining Workshops*, Brussels, Belgium, 2012, pp. 107–114.
[17] M. Kargar, A. An, and M. Zihayat, "Efficient bi-objective team formation in social networks," *Mach. Learn.*, vol. 7524, pp. 483–498, 2012.
[18] H. Yin, B. Cui, and Y. Huang, "Finding a wise group of experts in social networks," in *Proc. 7th Int. Conf. Adv. Data Mining Appl.*, Beijing, China, 2011, pp. 381–394.
[19] X. Wang, Z. Zhao, and W. Ng, "USTF: A unified system for team formation," *IEEE Trans. Big Data*, vol. 2, no. 1, pp. 70–84, Mar. 1, 2016.
[20] X. Wang, Z. Zhao, and W. Ng, *A Comparative Study of Team Formation in Social Networks*. New York, NY, USA: Springer, 2015, pp. 389–404.
[21] Z. Shi and F. Hao, "A strategy of multi-criteria decision-making task ranking in social-networks," *J. Supercomput.*, vol. 66, pp. 556–571, 2013.
[22] S. Rangapuram, T. Bühler, and M. Hein, "Towards realistic team formation in social networks based on densest subgraphs," in *Proc. 22nd Int. Conf. World Wide Web*, Rio de Janeiro, Brazil, 2013, pp. 1077–1088.
[23] L. Li, H. Tong, and N. Cao, "Replacing the irreplaceable: Fast algorithms for team member recommendation," in *Proc. Int. World Wide Web Conf. Committee*, 2015, pp. 636–646.
[24] K. Kamel, Z. Al Aghbari, and I. Kamel, "Realistic team formation using navigation and homophily," in *Proc. Int. Conf. Big Data Smart Comput.*, 2014, pp. 197–203.
[25] K. Kamel, N. Tubaiz, O. ALKoky, and Z. AlAghbari, "Toward forming an effective team using social network," in *Proc. Int. Conf. Innov. Inf. Technol.*, 2011, pp. 308–312.
[26] M. Juang, C. Huang, and J. Huang, "Efficient algorithms for team formation with a leader in social networks," *J. Supercomput.*, vol. 66, pp. 721–737, Mar. 23, 2013.
[27] L. Cheng and S. Man, "X2-Search: Contextual expert search in social networks," in *Proc. Conf. Technol. Appl. Artif. Intell.*, Taipei, Taiwan, 2013, pp. 176–181.
[28] M. Cheatham and K. Cleereman, "Application of social network analysis to collaborative team formation," in *Proc. Int. Symp. Collab. Technol. Syst.*, 2006, pp. 306–311.
[29] L. Cheng-Te and M. Shan, "Team formation for generalized tasks in expertise social networks," in *Proc. IEEE Int. Conf. Soc. Comput.*, 2010, pp. 9–16.
[30] R. D. Freeze, P. Lane, and S. Sasidharan, "Management teams: The case for balancing or unbalancing," in *Proc. 44th Hawaii Int. Conf. Syst. Sci.*, 2011, pp. 1–9.
[31] G. S. Easton and E. D. Rosenzweig, "Team leader experience in improvement teams: A social networks perspective," *J. Oper. Manage.*, vol. 37, pp. 13–30, 2015.
[32] Y. Liu, Q. Liu, R. Wu, E. Chen, Y. Su, and Z. Chen, "Collaborative learning team formation: A cognitive modeling perspective," in *Database Systems for Advanced Applications*. New York, NY, USA: Springer, 2016, pp. 383–400.
[33] H. Wi, J. Mun, S. Oh, and M. Jung, "Modeling and analysis of project team formation factors in a project-oriented virtual organization (ProVO)," *Expert Syst. Appl.*, vol. 36, pp. 5775–5783, 2009.
[34] R. S. Monclar, J. Oliveira, and M. Souza, "Analysis and balancing of social network to improve the knowledge flow on multidisciplinary teams," in *Proc. 13th Int. Conf. Comput. Supported Coop. Work Des.*, 2009, pp. 662–667.
[35] C. Dorn and S. Dustdar, "Composing near-optimal expert teams: A trade-off between skills and connectivity," *Lecture Notes Comput. Sci.*, vol. 6426, pp. 472–489, 2010.
[36] A. Farasat and A. G. Nikolaev, "Social structure optimization in team formation," *Comput. Oper. Res.*, vol. 74, pp. 127–142, 2016.
[37] I. H. Gutiérrez, C. A. Astudillo, P. Ballesteros-Pérez, D. Mora-Melià, and A. Candia-Véjar, "The multiple team formation problem using sociometry," *Comput. Oper. Res.*, vol. 75, pp. 150–162, 2016.
[38] W. Wang, J. Jiang, B. An, Y. Jiang, and B. Chen, "Toward efficient team formation for crowdsourcing in noncooperative social networks," *IEEE Trans. Cybernet.*, pp. 1–15, Sep. 2016.
[39] N. Dykhuis, P. Cohen, and Y.-H. Chang, "Simulating team formation in social networks," in *Proc. ASE/IEEE Int. Conf. Soc. Comput.*, 2013, pp. 244–253.
[40] B. Bulka, M. Gaston, and M. desJardins, "Local strategy learning in networked multi-agent team formation," *Auton. Agent Multi-Agent Syst.*, vol. 15, pp. 29–45, 2007.
[41] M. Zack, "Researching organizational systems using social network analysis," in *Proc. 33rd Annu. Hawaii Int. Conf. Syst. Sci.*, Maui, HI, USA, 2000, p. 7.
[42] H. Wi, S. Oh, and M. Jung, "Virtual organization for open innovation: Semantic web based inter-organizational team formation," *Expert Syst. Appl.*, vol. 38, pp. 8466–8475, 2011.
[43] C. Taramasco, J. Cointet, and C. Roth, "Academic team formation as evolving hypergraphs," *Scientometrics*, vol. 85, pp. 721–740, 2010.
[44] R. Michalski, S. Palus, and P. Kazienko, "Matching organizational structure and social network extracted from email communication," in *Business Information Systems*, vol. 87. New York, NY, USA: Springer, 2011, pp. 197–206.
[45] I. H. Jo, "The effect of social network diagrams on a virtual network of practice: a Korean case," *Asia Pacific Edu. Rev.*, vol. 10, pp. 525–534, 2009.
[46] M. H. Hahn, K. C. Lee, and D. S. Lee, "Network structure, organizational learning culture, and employee creativity in system integration companies: The mediating effects of exploitation and exploration," *Comput. Human Behav.*, vol. 42, pp. 167–175, 2013.
[47] O. Ardaiz-Villanueva, X. Nicuesa-Chacón, O. Brene-Artazcoz, M. L. Lizarraga, and M. T. Baquedano, "Evaluation of computer tools for idea generation and team formation in project-based learning," *Comput. Edu.*, vol. 56, pp. 700–711, 2011.
[48] M. Guijarro-Mata-García, M. Guijarro, and R. Fuentes-Fernández, "A clustering-based method for team formation in learning environments," in *Hybrid Artificial Intelligent Systems*. New York, NY, USA: Springer, 2016, pp. 475–486.
[49] H. Jaber, F. Marle, and M. Jankovic, "Improving collaborative decision making in new product development projects using clustering algorithms," *IEEE Trans. Eng. Manage.*, vol. 62, no. 4, pp. 475–483, Nov. 2015.
[50] Y. Yang and H. Hu, "Team formation with time limit in social networks," in *Proc. Int. Conf. Mechatronic Sci., Elect. Eng. Comput*, Shenyang, China, 2013, pp. 1590–1594.
[51] X. Han, Y. Liu, X. Guo, X. Wu, and X. Song, "Time constraint-based team formation in social networks," in *Proc. Int. Conf. Mechatronic Sci., Elect. Eng. Comput.*, Shenyang, China, 2013, pp. 1600–1604.
[52] M. E. Gatson and M. Desjardins, "The effect of network structure on dynamic team formation in multi-agent systems," *Comput. Intell.*, vol. 24, pp. 122–157, 2008.
[53] H. E. Aldrich and P. H. Kim, "Small worlds, infinite possibilities? How social networks affect entrepreneurial team formation and search," *Strat. Entrepreneurship J.*, vol. 1, pp. 147–165, 2007.
[54] L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. La, "Group formation in large social networks membership, growth, and evolution," in *Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, 2006, pp. 44–54.
[55] J. M. Alberola, E. Val, V. Sanchez-Anguix, A. Palomares, and M. D. Teruel, "An artificial intelligence tool for heterogeneous team formation in the classroom," *Knowl.-Based Syst.*, vol. 101, no. 1, pp. 1–14, 2016.
[56] DBLP, Dec. 22, 2016. [Online]. Available: http://dblp.uni-trier.de/
[57] "GAMS," Dec. 26, 2016. [Online]. Available: https://www.gams.com/

Authors' photographs and biographies not available at the time of publication.