

Predicting Teamwork Performance in Collaborative Project-Based Learning

Hoi Yan Lin¹ & Jia You¹

¹Department of Computing, The Hong Kong Polytechnic University, Hong Kong

Correspondence: Hoi Yan Lin, Department of Computing, The Hong Kong Polytechnic University, Hong Kong.
E-mail: 13901480r@connect.polyu.hk

Received: May 12, 2021 Accepted: June 20, 2021 Online Published: June 24, 2021

doi:10.5539/jel.v10n4p104 URL: <https://doi.org/10.5539/jel.v10n4p104>

Abstract

Pulse of the Profession, published by Project Management Institutes (2017), reported that failed projects always lacked (a) clearly defined objectives to measure progress and (b) poor communication between team members. Minimizing communication costs and maximizing trust levels are essential to improve the efficiency of team performance. This study's objectives required including how to formulate the problem and design the theoretical framework. The approach used involved a five-step team formation model with related definitions, including initial team forming, depending on group size, team agreement, role assignment, and team performance. The Predicting Teamwork Performance (PTPA) system was developed to help identify the functional roles of each member automatically. Role assignment provided a strong positive effect on team projects, while the role identification mechanism can assign team members responsibilities for some role(s) to enable learning. Self-assessment was used to identify team members' strengths and weaknesses so that team leaders could easily recognize suitable types of roles for each member. Three primary team performance indicators—"Good", "Pass" and "Marginal"—were reflected in the teamwork collaboration outcomes. The Predicting Teamwork Performance system reveals information about those outcomes through 1) individual performance indicator; 2) teamwork performance indicator; 3) personal skill sets results; 4) recommended skill sets improvements. The relationship between those indicators and practical roles was examined as analytical information for further project team formation.

Keywords: prediction teamwork performance, collaborative learning, project-based learning

1. Introduction and Background

Collaboration is an everyday activity in commercial enterprises for an individual to contribute knowledge and skills when qualifying for a new position. Good collaboration can foster project productivity. Collaborative learning can be part of team formation when a person has a task in mind and when he/she is willing to learn from others and contribute to forming a team to accomplish goal(s) effectively. Team projects are often essential for achieving project goals. For instance, a person collaborates with people in a virtual environment. The virtual team always happens in social dialogue. They aim to apply technological methods in real-life applications. GitHub is one example. The maximization of productive collaborations is still virtually unexplored, however. According to Abramowicz et al. (2003) collaboration can be more efficient than individual work by promoting creativity. Their analyses mentioned the significance of grouping people together into cooperative work teams.

1.1 The Interaction of Collaboration

Group support systems have the potential to encourage peers to construct or share knowledge and illustrate the importance of team development, such as in a group support system (GSS) that is a group of people who collaborate to work on complex decision making. The cooperative tasks mean participants agree on responsibilities and distribute those tasks across the group members who work independently until the assigned tasks have been completed.

The simple interaction of collaborative experiences is shown in Figure 1.

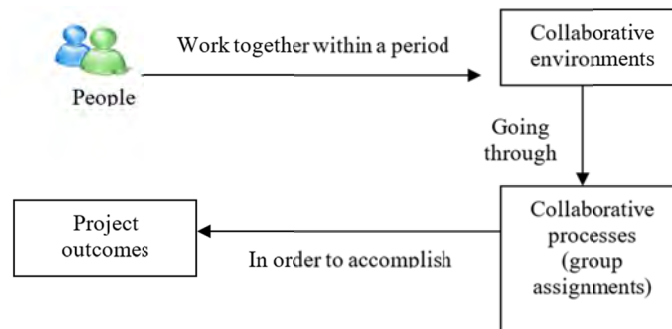


Figure 1. Interactions of collaborative experience

As Jain et al. (1996) stated, the question of how to maximize productive collaborations by manipulating role assignments and configuring teams is still largely unexplored. Therefore, this is a crucial consideration for team formulation. Wang and Zhang (2015) stated that the decision makers set sights on team formation that optimally determines their roles for a team project, team formation can be divided into a 5-step process.

A business journal article by Benoit (2012) mentioned an investigation by Pricewaterhouse Coopers about a review of company projects. They found 10,640 projects among 200 companies in 30 countries, in which “only 2.5% of the companies completed 100% of their IT projects”. Additionally, by looking at the *Harvard Business Review*, they illustrated that IT projects’ budgets tended to be 27% higher than the initial estimate. These statistics alone suggest that the cost of the project and the reason(s) for project failure need attention. According to Lappas et al. (2009), bottleneck edge cost is a cost function defining the minimum bottleneck cost of a pair of users. These cost models describe the properties of a team. A high collaboration cost will lead the team to reduce productivity to each member’s need to know, his or her role assignment, and the specification of team and negotiation instruments. In doing so, each team member needs to know their roles to make collaboration more productive. As a matter of fact:

- 1) Each team member encounters many unforeseen problems, such as misunderstanding the ideas between team members and unsatisfactory communication. Individuals gain knowledge and skills by investigating complex questions or exploring real-world problems and challenges (Hmelo-Silver, 2003)
- 2) How to create a new business model about how the team contributes is a challenge.

To resolve this issue, a teamwork formation model and a role assignment notion were developed to investigate task assignments in team projects. As a consequence, this study explores the following objective.

- Prediction of team members’ performance with task assignments in the project stage.

The first stage of a project starts with an initial team formation stage. The second stage of a project develops a team with a project review. The first phase occurs before the project starts, and the performance review is conducted in the final stage of the project.

1.2 The General Conditions for Teamwork and Teamwork Performance

Dafoulas and Macaulay (2001) indicated that role assignment and reassignment arise during collaboration in different stages of projects according to a group’s needs. A study by Stone and Veloso (1999), however, examined how role transfer can lead to formation change in a group of individuals working together. They applied their method to a robot soccer team and obtained an important result. The study was extended in Bowers et al. (2000) and Cannon-Bowers et al. (1997) focusing on team competencies to develop the five-step model, skills model, and trust direction algorithm. As a final point, the vital conditions of team performance are that all team members should have a clear position in a team, and their task assignments should not be ambiguous and conflict with the project goal. The success of task assignments would benefit team members and enable members’ to work together successfully.

2. Group Role Assignment (GRA) Problem

In this section, the mathematical model about the group role assignment problem (GRA) from H. Zhu et al. (2016, 2009) and M. Zhou (2006) is applied in this study. They identified how role negotiation, assignment and performance led to efficient collaboration among members. From H. Zhu et al. (2016, 2009), the 3-roles formation yields the best group performance. Thus, the assumption about group performance is expressed by the sum of individual performance values $\sum_{i=1}^T T$ that selects essential roles.

H. W. Kuhn (1995) solved the assignment problem in a polynomial. For example, a non-negative $m \times n$ matrix, where the i^{th} row and j^{th} column element represents qualification value of a role assigned the j^{th} job to the i^{th} member. Assignment of jobs to team members needs to be based on having maximum qualification values. The algorithm of bipartite graphs is $G = (S, T; E)$ with n member vertices (S) and n job vertices (T). Each edge has a non-negative value $v(i,j)$ to find the perfect matching with maximum qualification value. Let the function $y: (S \cup T) \rightarrow \mathbb{R}$ is a potential if $y(i) + y(j) \leq v(i,j)$ for each $i \in S, j \in T$. Then, the value of potential y is $\sum_{v \in S \cup T} y(v)$. Kuhn's method can find a perfect matching and a potential with an equal value that proves the optimality of both.

2.1 The Mathematical Models of Teamwork Performance

2.1.1 Problem Definitions

Definition 0 (Identification of project): $P ::= \langle Sstage, Ot \rangle$, where $Sstage$ is the project precedence relationship and is sequential, which means that project decomposes into smaller sequential phases. Ot is the overall time the project took to finish. In general, the potential sequent phases, n , then, $\sum_{i=0}^{n-1} (n - 1_i)$ is a possible option to complete the project.

Definition 1 (Identification of Object): $O ::= \langle Oid, Os \rangle$, where, Oid identifies the object, Os is a data structure, those values are attributes, the properties of objects stated in Table 1.

Definition 2 (Group Size and Capability): Non-negative integers $|Gs|$ express a group size, Gs and $|Ga|$ explicates the group ability of the role set Ga, i, i_1, i_2, \dots that denotes the group characteristics like j, j_1, j_2, \dots to j_n .

Definition 3 (Formation with group characteristics): A qualification matrix Q (refer to training interest in Table 1) is an $m \times n$ matrix, where $Q[i,j] \in [0,1]$ expresses the value of role $i \in N (0 \leq i \leq m)$. For the group of characteristics defined as $j \in N (0 \leq j \leq n)$, $Q[i,j] = 0$ indicates the lowest value and 1 indicates the highest. Note: N is non-negative integers.

Definition 4 (Role Assignment): A role assignment matrix RA , which is defined as an $(m \times n)$ matrix, where $RA[i,j] \in [0,1] (0 \leq i \leq m, 0 \leq j \leq n)$, denotes whether role i is assigned to the characteristics of group j or not. $RA [i,j]=1$ means assigned and 0 means not assigned (refer to Table 2).

Definition 5 (Task Assignment): Given a task $T = \{t_1, t_2, t_3, \dots, t_n\}$ and each of role $R \in T_n$. Each role is assigned the tasks in each stage of the project (refer to assumption 4). The maximum number of tasks assigned to R and all Ts are assigned to roles like $T_n > R$. The same sets of roles persist in the stages of the project, and task assigning is step-3 in the team formation model.

Definition 6 (Task Performance): Minimizing underperformance is necessary in the project stages. The problem can be stated as X , which minimizes $\sum_{i=1}^T f_i(x_i, y_i)$ subject to the constraints;

$$\begin{aligned} t_i(x_i, y_i) + y_i &= y_{i+1}, & i &= 1, 2, \dots, n \\ g_i(x_j) &\leq 0, & j &= 1, 2, \dots, n \\ h_i(y_j) &\leq 0, & k &= 1, 2, \dots, n \end{aligned}$$

where x_i is the i^{th} control variable, y_i is the i^{th} control variable such as the group size and task interdependence. Also, both x_i and y_i are control variables, and f_i is the contribution of the i^{th} stage to the total objective functions; t_i, g_j and h_i are functions of x_j, y_k and x_i and y_i , respectively, and n is the number of stages.

Definition 7 (Re-assignment): A team member continues to improve from training or feedback. Os is a data structure (defined in definition 1); T_{nw} is the new task assignment after ongoing improvement from training. So, it will have $Ga, Os \in T_{nw}$. The new assignment, as $\sum_{i=1}^T Tnw$, means that every role has some new tasks within the group.

Definition 8 (Determined Initial Stage Performance): A group is defined as the sum of the assigned roles' value as $F(o) = \sum_{j=1}^m \sum_{i=1}^n (Wm * Xm) * (Wn * Xn)$, where W is weight and X is input as in Table 1. The weight Wm is a coefficient that determines the strength values between two nodes. The result from the multiplication is summed and passes through a transfer function that maps the inputs into the range of 0 to 1. In such, maximization of $W(m*n)$ is needed and the problem of team performance is $\max W (mn_1, mn_2, mn_3, \dots, mn_n)$ subject to $\sum_{k=1}^{mn} Wnm \leq 1$ So, $X_i \geq 0 \forall i \in \{1, 2, 3, \dots, n\}$.

Definition 9 (Group Performance Review): Individual performance σ_0 is defined as the sum of the task performances $\sigma_0 = \sum_{i=1}^T \sum_{j=1}^{Tnw}$. So, group performance is $Z = \sigma_0 * \sum_{k=1}^{mn} Wnm$.

Teamwork performance may be adjusted by the team formation.

For instance,

Assume the total number of project stages is 6	The group performance $Z = \sigma_0 * \sum_{k=1}^{mn} Wnm$	The possible combination of team members corresponding to the group performance indicator (assume three members with independent tasks assigned in each stage)
1	91.3/100	A B C
2	92.2/100	B C A
3	93.0/100	C B A
4	93.5/100	C A B
5	92.5/100	B A C
6	91.4/100	A C B

From the above example, the best performance indicator is 93.5. Team performance is expressed by the sum of individual performance values that is defined in definition 10.

Definition 10 (Determined Next Stage Performance): The next stage performance σ_1 is defined as the sum of the task performance $\sigma_1 = \sum_{i=1}^T \sum_{j=1}^{Tnw}$. So, the next stage team performance is

$$Z\sigma = Z + \sigma_1 * \sum_{k=1}^{mn} Wnm.$$

2.1.2 Theoretical Framework: The Five-Step Team Formation

According to Williams et al. (1991) theory, an individual achieving goals that are related to a team project's timely completion and quality is likely unfair. Individual assessments should correspond to each member's minutes spent and materials produced. The proposed five-step team formation model references GRA, and the following shows the theoretical steps.

Step 1—Initialization

The group size and capability with non-negative integers |Gs| and |Ga| as its input and having to do with a qualification matrix Q to express the role $i \in N (0 \leq i \leq m)$.

Step 2—Role Assignment

A role assignment matrix RA, which is defined as an $m \times n$ matrix, where $RA[i,j] \in [0,1] (0 \leq i \leq m, 0 \leq j \leq n)$, denotes whether role i is assigned to the group j or not. $RA [i,j]=1$ means assigned and 0 means not assigned (refer to Table 2). The same sets of roles persist in the stages of the project.

Step 3—Task Assignment

Given a task $T = \{t_1, t_2, t_3, \dots, t_n\}$ and each of role $R \in T_n$. Each role is assigned a task in each stage of the project (refer to assumption 4). The maximum number of tasks assigned to R and all Ts are assigned to roles like $T_n > R$. The same sets of roles persist in the project stages, and task assigning is step-3 in the team formation model.

Step 4—Task Performance

Underperformance is minimized in the project's stages. The problem can be stated as X, which minimizes $\sum_{i=1}^T fi(xi, yi)$ subject to the constraints in definition 6. In the direction of the communication cost algorithm, it can find the minimum communication cost and discover the probability of the communication node so that finding the trusted ratio is according to the trusted direction algorithm and estimate of task performance.

Step 5—Re-Assignment

A new assignment, as $\sum_{i=1}^T Tnw$, is made based on higher communication cost and performance review after the initial stages (Refer to definition 8 and 9). This means that every role must have some new tasks within the team. For that reason, the team should be more productive than they would normally be in a typical environment, so the process returns to definition 7.

Step 6—End of Stage

Definition 8 to 10 stated that task performances are highly dependent on the definition of 4 to 8 various conditions. The five-step team formation model repeats steps 2 to 5 until no further improvement can be found, and then it returns the result.

Table 1. The training interests (Qualification Values = Q)

Member	Position - Leader	Position - Team Member (Consolidation/Report Writer)	Position - Team Member (Information Seeker/Program Design)	Position - Team Member (Program Design/Development)
A	0.33	0.67	0.51	0.86
B	0.02	0.14	0.71	0.73
C	0.22	0.79	0.04	0.53

Table 2. The initial training plan (Role assignment = RA)

Member	Position - Leader	Position - Team Member (Consolidation/Report Writer)	Position - Team Member (Information Seeker/Program Design)	Position - Team Member (Program Design/Development)
A	1	0	0	1
B	0	0	1	0
C	0	1	0	0

Table 3 shows the input attributes among the general project tasks and expected output

Table 3. The descriptions of the input attributes, weighting, skill sets, roles, and indicators

Input Parameters	Weighting	Skill Sets	Role Assignment	Prediction Performance Indicator
Good at oral & written communication skills and proofreading skills	Poor	English	Action	Good
Imagination and creative reasoning	Satisfactory	Interpersonal/	Task	Pass
Able to work with others	Good	Communication	Social	Marginal
Continuously update personal skills and knowledge	Very Good	Leadership		
Able to accept responsibility	Excellent	Creativity		
Effective and on time contributions		Technologically		
Able to understand and solve complex problems		Savvy		
Think flexibly and have a personal sense of challenge (i.e. having new perspectives on problems)		Logical		
Organizing knowledge about analytical problems & solving using essential techniques				
Good at thinking logically and analytically in the problem solving environment				
Able to work independently				
Good at controlling concurrent problems and solutions				
Proposed methodology to achieve the project specifications				
Accepting criticism gracefully				

Following the above theoretical framework, the individual performance data are collected from the academic results, past project scores and rubric sets on the Blackboard. After collecting all the data, a classification of data is required for developing the training set. It then repeats the classification using cross-validation if there is high over-fitting. In the context of definition 8, a group defined the sum of the assigned roles' qualifications as $F(o) = \sum_{j=1}^m \sum_{i=1}^n (Wm * Xm) * (Wn * Xn)$, where W is weight and X is as entered in Table 1.

2.1.3 Five-Step Modelling and Assumptions

The following are the assumptions of the proposed model:

- 1) Each member is represented as a node
- 2) Each member is assigned tasks before starting the project.
 - 2.1) Each member is assigned at least one task.
 - 2.2) Each project requires a set of skills.
- 3) There is limited time for each project.

4) A team allocates tasks as E_i at the beginning of the project.

4.1) The contribution of the team member to team $C(E_i)$ is assumed to be adjusted at the middle stage of a project, which is represented by binary E_i -vector, $x = (x_1, x_2, x_3, \dots, x_n)$, in which $x_1 = 1$ means that each member has a minimum of one task. The example is like 1 = task assigned, 0 = no task assigned,

Member	1	2	3	4	5	6
Task x=	1	1	1	1	1	1

1st stage

Member	1	2	3	4	5	6
*RA x=	0	1	1	0	1	1

2nd stage
According to the performance of 1st stage
*refer to definition 7

5) The levels of task interdependence in the stages of the project are

Mutual (highly interdependent): $1 \leftrightarrow 2 \leftrightarrow 3$ (task interdependence within the team)

Sequential (medium interdependence): $1 \rightarrow 2 \rightarrow 3$ (linked task within the team)

Collection (lower interdependence): Resource to $\begin{matrix} 1 \\ \updownarrow \\ 2 \end{matrix} \rightarrow 3$ (collective tasks within the team)

5.1) The minimum contributions of a team member are shown by $\min(x_n) = (0 \leq \min(x_n) \leq 1)$ in each stage of the project.

5.2) The maximum contributions of a team member are shown by $\max(x_n) = (0 \leq \max(x_n) \leq 1)$ in each stage of the project.

5.3) A local maximum x^* like the interval of maximum contributions is $(x_n) = (0 \leq \max(x_n) \leq 1)$ of the whole team members whose performance is greater than or equal to all its neighbors.

6) Assume assigned tasks are static in the 1st stage.

7) Assume α is positive communication between a pair of members

7.1) if α is close to 0, the original team formation structure will be mostly preserved

7.2) if α is close to 1, a new target member will be constructed.

8) All nodes must know their role assignment within the team.

9) No negative communication paths.

10) A “Free-rider” may be considered within the team.

With the above assumptions, the proposed five-step team formation model becomes

1) **Forming.** People discuss the project. The group tries to reach agreement(s) and then consolidate the agreements. For example, a group has three members (refer to Table 1), which can be expressed as $O_1 \dots O_3$. Each group has an identification as Oid , and O_s that is a data structure; those values are attributes $O ::= \langle Oid, O_s \rangle$, the properties of the groups are stated in Table 1. Members start to rely on collaboration such that each member is doing only one task $|T| \geq 1$.

2) **Storming.** A member needs to have a sensitive response to the burdens of the project after consolidating the agreements. On the topic of James et al.’s (1993) within-group agreement, $r_{WG(j)}$ was toward the group variance in scale scores like $s_{x_j}^2$, which means that $r_{WG} = 1 - (s_{x_j}^2 / \sigma_E^2)$. If an individual does not know exactly how he/she is expected to perform due to unclear and abstractly-set expectations, there is intragroup conflict. The corresponding situation in Table 1, B and C, have role conflict like both dealing with program design, and assumed $T \approx T_{nw}$. A conflict is defined as $m \times n$ matrix A ($A[i_1, i_2] \in \{0, 1\}, 0 \leq i_1, i_2 < m$), where $A[i_1, i_2] = 1$ expresses that $C i_1$ conflicts with $B i_2$. To solve this intragroup conflict, the number of members must meet the requirement of $|T| \geq \sum_{j=0}^{n-1} T[j]$.

3) **Assigning.** Members need to exchange relevant information during the project period. Then, $E_t \leq 1$ and the requirements of the members are defined as $\sum_{j=0}^{n-1} T[j]$. A member needs to improve the technical or academic skills as “S” in responses from teachers as “F”. So, the new requirements of the members can be defined as $\sum_{j=0}^{n-1} T(S + F)$. It follows that the team members start to develop group cohesion.

- 4) **Performing.** Each member needs to collaborate more productively than if they are working in a typical environment. For example, in Table 1, 3-person team, the contributions of members A, B, and C as (C_{aa}, C_{ab}, C_{ac}) or (C_{ba}, C_{bb}, C_{bc}) or (C_{ca}, C_{cb}, C_{cc}) , the sum of all members is 1.0. If member “A” claims that he/she has a high contribution (i.e. $C_{aa} = 0.8$), the total contributions of other members is correspondingly lower (i.e. $C_{ab} + C_{ac} = 0.2$). Hence, the productivity is lower in the whole team.

The incentive design from Chiu et al., (2020) extends the free-rider concept to find the minimum contribution of each team member. For example, as a 2-person team, member A’s contribution is G_A , the minimum contribution of $G_A = \min [(D_A / ES) G, \alpha G \text{ (Note 1), } 100]$, where $D_A = S_A / (S_A + S_B)$ = member A’s contribution. Hence, share S_n may not sum to 1.0; ES = equal share = 0.5 for two members in the team; α = preset scalar ≥ 1 ; and G = overall project score. Thus, (D_A/ES) is an estimated extent of member A’s contribution to G relative to equal share ES. With the intention of D_A and G_A , three common situations in team projects are $D_A > ES, G_A > G, D_A = 0$ and $G_A = 0$ and $D_A = ES$ and $G_A = G$. Thus, minimizing intragroup conflict as minimum $(\sum_{j=0}^{n-1} T[j])$ is possible as each member needs to have the surfacing solutions in the project activity.

- 5) **Adjourning.** All members retain responsibilities and rights in project activity. The calculation of the initial stage performance to find the weight of the team performance is based on definition 8, which can be calculated as $F(o) = \sum_{j=1}^m \sum_{i=1}^n (Wm * Xm) * (Wn * Xn)$, where W is weight and X is input (refer to Table 1). After establishing the initial stage performance, the group performance is defined in definition 9 as $Z = \sigma_0 * \sum_{k=1}^{mn} Wnm$. All members accept the need to collaborate on tasks to finish the project with good results.

3. Operational Procedures and PTPA Framework

To achieve the accuracy of predictive model, this study used students’ academic records during the pre-processing stage to extract features like English skills, leadership skills, communication skills, technological savvy skills, logical skills, and hardware skills. The proposed framework for measuring and predicting teamwork’s performance was developed. The training set of both source and target output needed to have pre-processing processes such as cleansing, features encoding, and normalization of academic results. With these features, clustering and then labelling the data are necessary to extract the strongest correlated indicators. The data with the strongest correlated indicators were used in different classification algorithms and then in finding the most accurate model for predicting student performance in the project-based assignment. The figure shown in Figure 2 was developed to measure and predict the teamwork's performance.

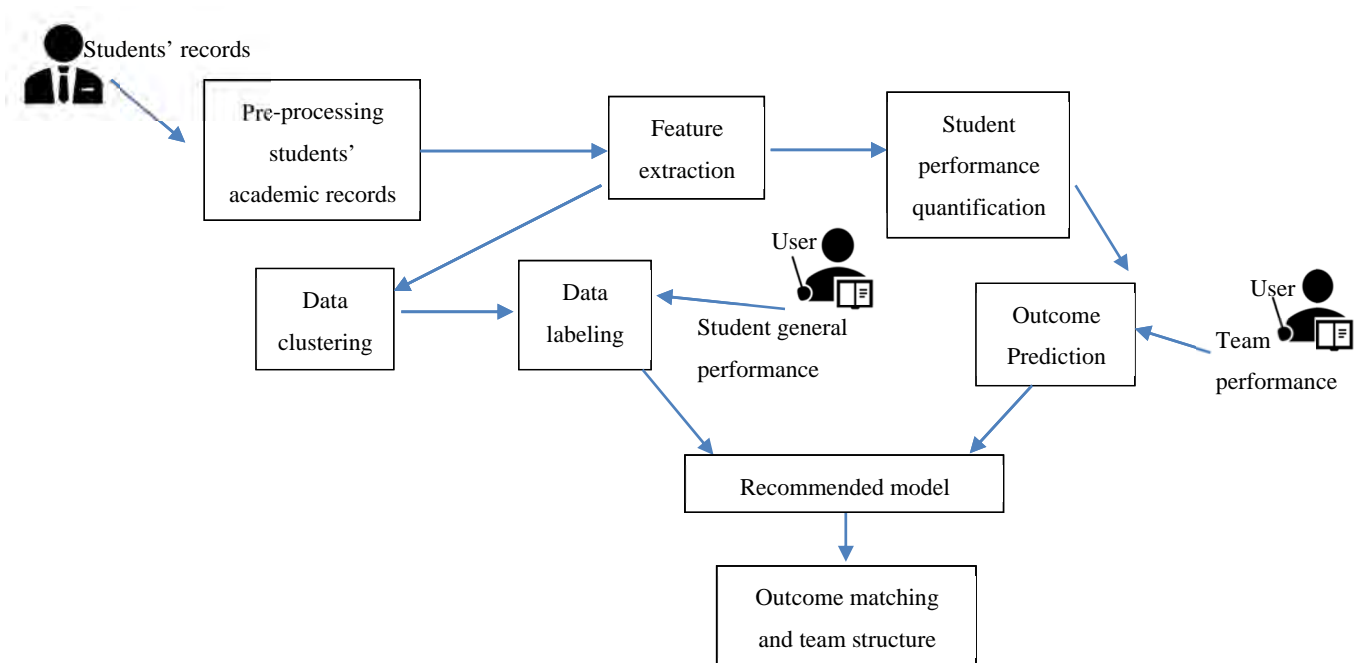


Figure 2. The framework for measuring and predicting teamwork’s performance

Nearly 851 records were collected from three-team projects (each from one undergraduate program) related to programming/technological, hardware development and generic IT study throughout three academic years. Then, the subjects were mapped to the relevant skills as features for training the machine learning model. From those records, two datasets were used for model formation and model analysis to answer the research questions. In order to achieve the highest accuracy machine learning model in our dataset, we divided it into the training set and testing set to build up a model. Two-third of them were the training set, and the rest of them were the testing set. The testing set was to validate the model building, and data points in the training set are excluded from the testing set.

3.1 Data Handling Context

The data handling and data collection sources from 16/17, 17/18, 18/19 and 19/20 academic years.

Table 4. Data summary

Academic Year/Subject code and name	2016/2017	2017/2018	2018/2019	2019/2020
EIE3120 - Network Technologies and Security	--	--	38	39
EIE3360 - Integrated Project	40	33	34	33
EIE3105 - Integrated Project	--	86	88	86
EIE3106 - Integrated Project	--	60	57	71
EIE2264 - Computer Programming	--	78	71	78

According to the categories' skill set, labelling data for each task was done, and some complex data were found. The data from the assessment form were input to separate excel files. Each record contains all members of a team and the exact score of the whole team. The six categories (including English, communication, creativity, leadership, programming, logical) were used as the inputs for each member and each is represented by "0" or "1" by calculating the workload areas satisfying at least half in the corresponding category. "0" represents that the member does not have the skill, "1" represents that the member has the skill. The output is the underperformance, which is represented by "0" or "1". "0" represents that the member does not underperform, "1" represents that the member underperforms. Normalization was employed to calculate the sub-scores to determine underperformance and in pre-processing the dataset; for example, the normalized project report = project report score/full project report score. The peer assessment score divided by the total peer assessment score normalizes the peer assessment.

3.2 Role Descriptions

"Belbin's Team Roles" show that THREE roles achieve maximization of productive collaborations in universal, technical, hardware and programming projects. From Gareth Bell (2013), Belbin claimed that teams that accomplished their goals regularly mostly have a healthy mix of the various team roles. Therefore, a "Task Role" for members is to take on helpful information-seeking and enjoy being procedural technicians like recorder and consolidator, or report writer or presenter. The "Social role" is the person who takes a leadership role within the team, and that person can lead the group's day-to-day activities. Every member on the team respects and listens to the gatekeeper, observer, commentator, and harmonizer. The "Action role" is the last role suggested for members who finish the tasks in time. For example, members get the right to work, such as extracting the best possible solution from different sources, preparing the detail-oriented tasks, solving conflict among team members, ensuring thorough completion, and putting ideas into action(s). To deal with the role categories above, the skill mapping sets were designed below.

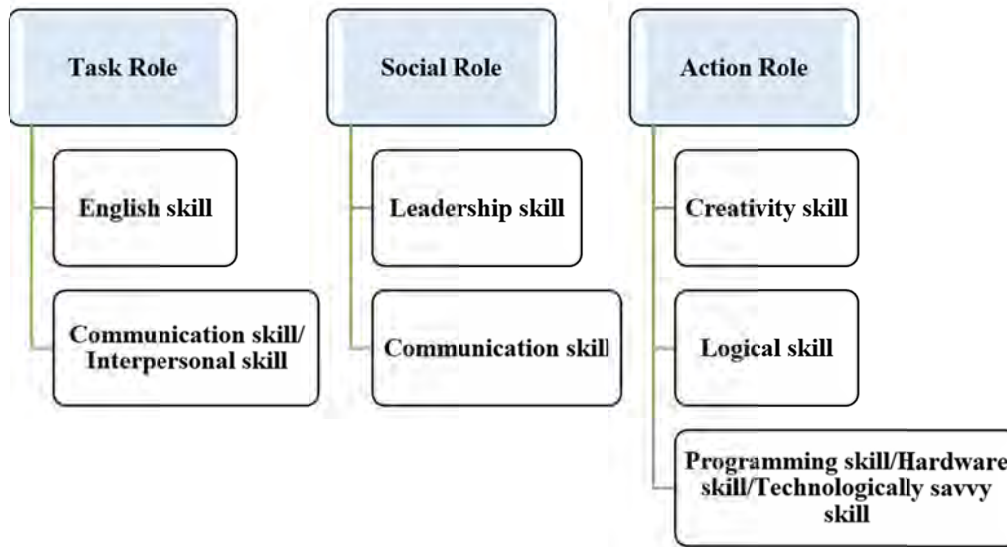


Figure 3. Different skill sets associated with roles

Within the context of the skill sets to assign roles, students are qualified by their academic results to assign the workload areas. They categorized data into the relation of each category and skill; that is, which belongs to which category. For example, skills of “English”, “Creativity”, “Programming”, “Leadership”, “Communication”, “Logical” could be assigned to 14 different tasks listed in Figure 4.

Area ID	Area Description	Skill ID	Skill Categories	Area ID	Role	Skill ID
1	'System design and planning',	1	English skill	2,7	Action Role	4,5,6
2	'Data collection and analysis',	2	Communication skill	7,8,14		
3	'System development',	3	Leadership skill	6,10,11,12,13,14		
4	Thinking logically and analytically in the problem-solving environment',	4	Creativity skill	5,9		
5	'Graphics design',	5	Programming skill	1,3		
6	'Time management skill (e.g. team phase completion date, etc.)',	6	Logical skill	1,2,4		
7	Good at oral & written communication skills and proofreading skills',				Task Role	2,3
8	'Accepting criticism gracefully',				Social Role	1,2
9	Imagination and good at creative reasoning',					
10	'Achieve team objectives',					
11	'Able to direct the work of others',					
12	'Able to work independently and direct the work of others',					
13	Proposed methodology to achieve the project specifications',					
14	'Project management and coordination',					

Figure 4. Role mapping using project tasks and skill sets

The calculation formula checks whether the sum of areas in a category is greater than or equal to half of that category’s total areas. The result is 1 when it is true; otherwise, it is 0. For example, if a team member selected area 1, 2, 3, the weighting of each area and role calculation is shown in Figure 5.

Area ID	Weighting	Score	Calculation	Score
1	Good	60	(20) / 2	10
2	Poor	20	0 / 3	0
3	Satisfied	40	0 / 6	0
			0 / 2	0
			(60+40) / 2	50
			(60 + 20) / 3	26.6666667
Action Role			50 + 26.6666667	76.6666667
Task Role			0	0
Social Role			10	10

Figure 5. Example for computation of the role

Referring to Figure 4 for computing role assignment, the weightings of “Excellent”, “Very Good”, “Good”, “Satisfactory” and “Poor” were 100, 80, 60, 40 and 20 respectively. For instance, for the English skill mapped to area ID 2 and 7, the score for English skill is 20/2 (total among English areas is 2) = 10. The highest score is “Programming skill” for that member because a member only has programming skill and logical skill provided and the total score is 76.67. Therefore, he/she is assigned an “Action Role” (the action role included skill ID 4, 5 and 6).

3.3 Environment Setup

To utilize the provisional dataset, the attributes’ values were changed and normalized into ranges [0,1] with a scale [1-5] before feeding them into the classification models. The “Anaconda” integrated development platform, which was used to visualize the dataset and analyze the attributes, was also used to train the machine learning model with different algorithms and evaluate them to select the most accurate algorithms. The Python language, Jupyter notebook application, and sci-kit learn library were installed for data analysis and visualization. Sharma et al., (2015) declared that Linear Discriminant Analysis (LDA) is efficient in dealing with small datasets. Additionally, the Multilayer Perceptron (MLP) algorithm, the Support Vector Machine (SVM) algorithm, Naïve Bayes (NB) and K Nearest Neighbour (KNN) algorithms were considered to train machine learning model. The teamwork-performance prediction system (PTPA) was developed to deploy the model and and display the teamwork performance with performance indicators like Good, Pass and Marginal. Also, the PTPA was behind the web application firewall, a specific application firewall that prevents exploiting web vulnerabilities. The system runs on the virtual machine and is open for clients from the internet and campus network.

3.3.1 High-Performance Computing Platform Structure

The web interface distributes the project tasks for the team project and collects the data for model creation. The PHP Laravel framework is used as the front-end portal to manage the team members’ responses, which is part of the input of the teamwork-performance prediction system. The web application runs on a Virtual Machine hosted by University Information Technology Services Centre. The server was created using docker and Nginx. Dockerize is a web application to make sure the web application could run anywhere. Then, docker and docker-compose were installed on the Virtual Machine hosted by University Information Technology Services Centre, and the web application was deployed using git.

3.4 The Teamwork Performance Prediction System

The visual composition was separated into user view and administrator view. The objective was to enhance efficiency and perform essential functions between students and teachers. In the user view, each team member needs to complete two surveys. The first survey is self-assessment (see Figure 6), which encourages the team members to examine their skills. This process assists in project role assignment.

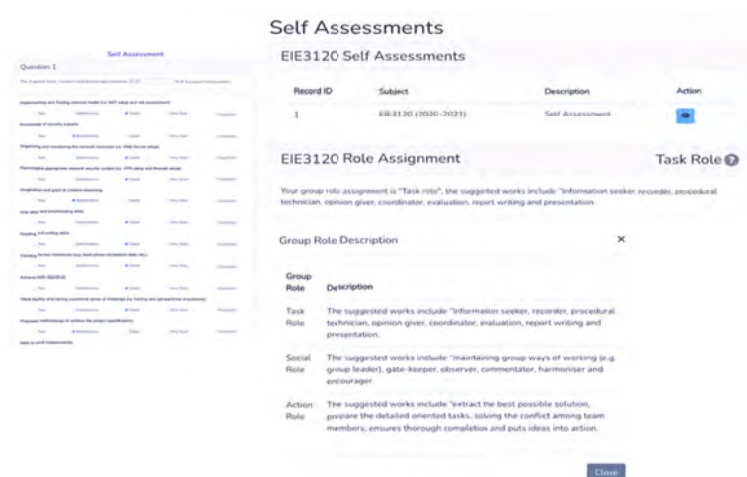


Figure 6. The self-assessment layout of the teamwork-performance prediction system

The peer assessment survey needs to be completed after the project ends. An optional feature of PTPA allows team members to form groups themselves, and to choose their group members after clicking the confirm button

(see Figure 7). The peer assessment survey provides a structured learning process for teammates to offer feedback to each other on their work, by calculating the total workload distribution provided to expose the free rider in the team project (see Figure 8). Moreover, the predicted performance indicator in group and individuals would be displayed in table format. The skill categories' evaluation results are also assessed by their teammates (see Figure 9).

The recommendation skill sets illustration for different natures of the team project is displayed. In Figure 9, recommended improvements are provided for use if the rating of a specific skill is below "Good". The purpose of this arrangement is to help underperforming teammates follow the recommended improvements to improve their skills.



Figure 7. Forming group layout of the teamwork-performance prediction system

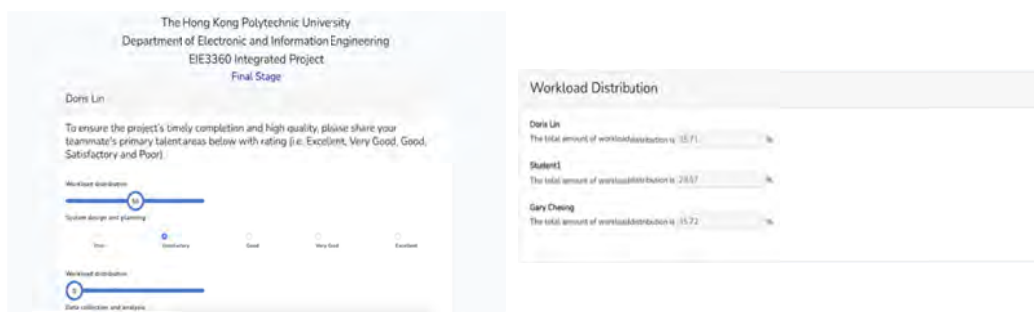


Figure 8. The final assessment layout of the teamwork-performance prediction system

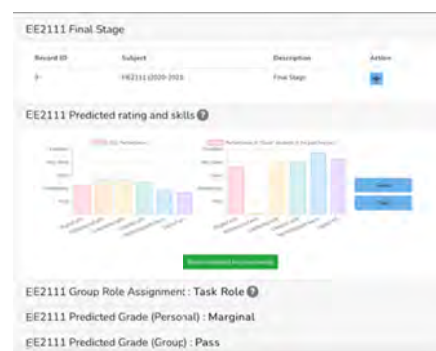


Figure 9. The prediction and assessment results of the teamwork-performance prediction system

4. Research Findings

Descriptive statistics are provided in this study. The online user acceptance survey was administered to participating students at the end of the semester. The survey asked students to check the box that best described their observations, the notation being Strongly disagree = 1, Disagree = 2 (tend to disagree more than not),

Disagree = 3, Agree = 4, Agree = 5 (tend to be more agreeable), Strongly agree = 6 and Not applicable = 0. The 19 questions comprised four aspects: learning patterns, teamwork cooperation, learning motivation, and skill diversity to achieve the learning outcomes for their programmes. Over 50% of students recognized that the PTPA system helped them complete the project learning outcomes and agreed that the PTPA system improved their learning patterns and understanding of their strengths and weaknesses (see Figure 10).

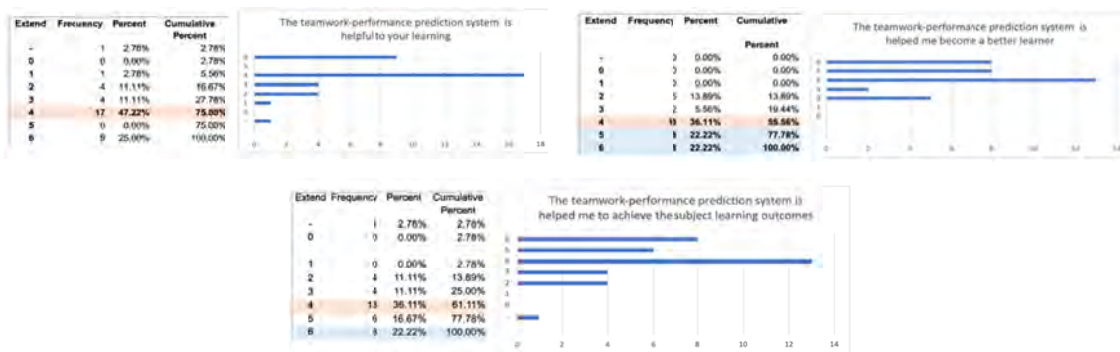


Figure 10. The survey results about the project learning outcomes

The teamwork cooperation element is one of the significant components. Many students actively use the PTPA system to accommodate their team members and understand the project workload distributions. For example, around 60% of students agreed that the PTPA system helped them cooperate with team members effectively and distributed the workload appropriately (see Figure 11).

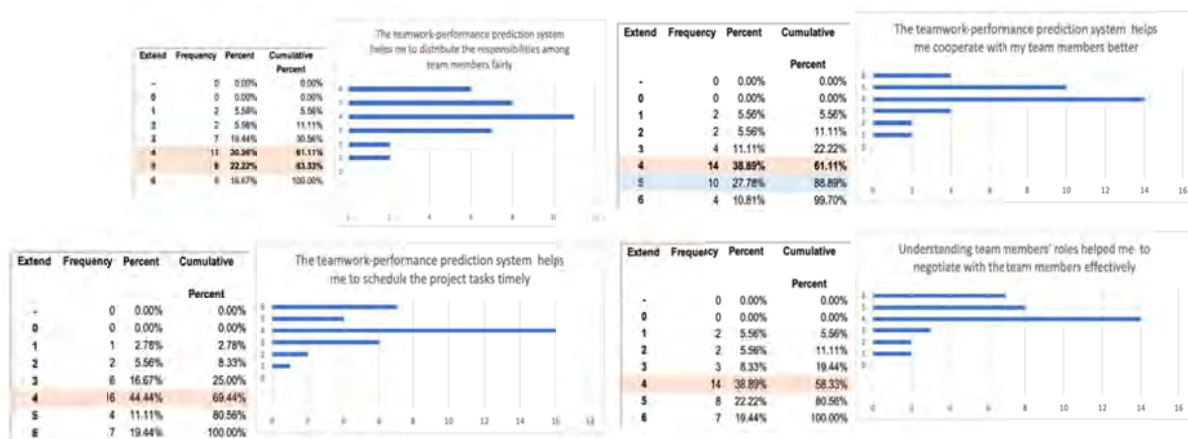


Figure 11. The survey results about the teamwork cooperation

More than 50% of students admitted the PTPA system helped them understand the project requirements and motivated them to learn further. Also, many students started to visualize their future development plan (see Figure 12).

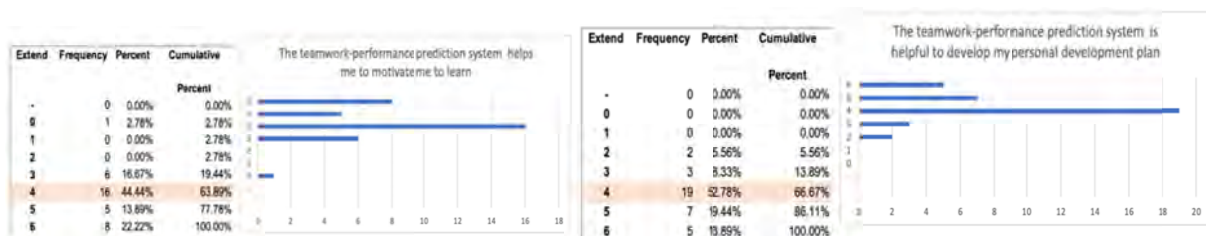


Figure 12. The survey results about the learning motivation

5. Conclusion

A five-step team formation model was developed to analyze teamwork performance. Referring to the preliminary analysis, the results show that the model is adequate and the relationship between the procedures of PTPA system, skills and role assignments can resolve team conflicts and develop reliable relationships between teamwork and teamwork formation. This study highlights vital aspects of successful project management.

Acknowledgements

This research project is funded by an internal teaching grant “Teaching Development Grant” from The Hong Kong Polytechnic University.

References

- Abramowicz, W., Kaczmarek, T., & Kowalkiewicz, M. (2003). *E-Collaboration System Designed to Improve Learning Processes*. Informing Science and IT Education Conference in Pori, Finland. <https://doi.org/10.28945/2664>
- Benoit, H.-V. (2012). The Cost of Bad Project Management. *Business Journal*.
- Bowers, C. A., Pharmed, J. A., & Salas. (2000). E. When member homogeneity is needed in work teams: A meta-analysis. *Small Group Research*, 31(3), 305–327. <https://doi.org/10.1177/104649640003100303>
- Cannon-Bowers, J. A., & Salas, E. (1997). A framework for developing team performance measures in training. In M. T. Brannick, E. Salas & C. Prince (Eds.), *Team Performance Assessment and Measurement – Theory, Methods, and Applications*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Chiu, M. M., Woo, C. K., Shiu, A., Liu, Y., & Luo, B. X. (2020). Reducing costly free-rider effects via OASIS. *International Journal of Comparative Education and Development*, 22(1), 30–48. <https://doi.org/10.1108/IJCED-07-2019-0041>
- Dafoulas, G. A., & Macaulay, L. A. (2001). Macaulay. Facilitating group formation and role allocation in software engineering groups, in Proc. *ACS/IEEE Int. Conf. Comput. Syst. Appl.*, 352–359. <https://doi.org/10.1109/AICCSA.2001.934012>
- Gareth, B. I. (2013). Teamwork makes the teamwork: An interview with Dr Meredith Belbin. *Human Resource Management International Digest*, 21(2), 45–47. <https://doi.org/10.1108/09670731311306850>
- Henry, S. M., & Stevens, K. T. (1999). Using Belbins’ leadership role to improve team effectiveness: An empirical investigation. *The Journal of Systems and Software*, 44, 241–250. [https://doi.org/10.1016/S0164-1212\(98\)10060-2](https://doi.org/10.1016/S0164-1212(98)10060-2)
- Hmelo-Silver, C. E. (2003). Analyzing collaborative knowledge construction: Multiple methods for integrated understanding. *Computers and Education*, 41(4), 397–420. <https://doi.org/10.1016/j.compedu.2003.07.001>
- Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, 29(3), 31–44. <https://doi.org/10.1109/2.485891>
- James, L. R., Demaree, R. G., & Wolf, G. (1993). r_{wg} : An assessment of within-group interrater agreement. *Journal of Applied Psychology*, 78(2), 306–309. <https://doi.org/10.1037/0021-9010.78.2.306>
- Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1–2), 83–97. <https://doi.org/10.1002/nav.3800020109>
- Lappas, T. et al. (2009). *Finding a team of experts in social networks*. In KDD ’09. <https://doi.org/10.1145/1557019.1557074>
- Sharma, A., & Paliwal, K. K. (2015). Linear discriminant analysis for the small sample size problem: an overview. *Int. J. Mach. Learn. & Cyber.*, 6, 443–454. <https://doi.org/10.1007/s13042-013-0226-9>
- Stone, P., & Veloso, M. (1999). Task decomposition, dynamic role assignment, and low-bandwidth communication for real-time strategic teamwork. *Artificial Intelligence*, 110, 241–273. [https://doi.org/10.1016/S0004-3702\(99\)00025-9](https://doi.org/10.1016/S0004-3702(99)00025-9)
- Wang, J., & Zhang, J. (2015). A win-win team formation problem based on the negotiation. *Eng Appl Artif Intell.*, 44, 137–152. <https://doi.org/10.1016/j.engappai.2015.06.001>
- Williams, D. L., John, D. B., & Jone, R. (1991). Team projects: Achieving their full potential. *Journal of Marketing Education*, 13(summer), 45–53. <https://doi.org/10.1177/027347539101300208>
- Zhu, H. (2016). Avoiding conflicts by group role assignment. *IEEE Trans. on Systems, Man, and Cybernetics:*

Systems, 46(4), 535–547. <https://doi.org/10.1109/TSMC.2015.2438690>

Zhu, H., Liu, D., Zhang, S., Zhu, Y., Teng, L., & Teng, S. (2016). Solving the many to many assignment problems by improving the Kuhn-Munkres algorithm with backtracking. *Theoretical Computer Science*, 618, 30–41. <https://doi.org/10.1016/j.tcs.2016.01.002>

Zhu, H., & Zhou, M. C. (2006). Role-based collaboration and its kernel mechanisms. *IEEE Trans. on Systems (Man and Cybernetics, Part C)*, 36(4), 578–589. <https://doi.org/10.1109/TSMCC.2006.875726>

Zhu, H., & Zhou, M. C. (2009). M–M role-transfer problems and their solutions. *IEEE Trans. Syst. Man, Cybern. A. Syst., Humans*, 39(2), 448–459. <https://doi.org/10.1109/TSMCA.2008.2009924>

Notes

Note 1. GA is covered at αG .

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).