Meta-heuristic approaches to tackle Skill Based Group allocation of Students in Project Based Learning Courses

Ravneil Nand * ‡, Anuraganand Sharma * §
* School of Computing, Information and Mathematical Sciences
The University of the South Pacific, Suva, Fiji. *http://scims.fste.usp.ac.fj/*

‡ nand_ra@usp.ac.fj, § sharma_au@usp.ac.fj

***Abstract*—In the arena of software engineering, Project Based Learning (PBL) is one of the fundamental components of practical based assessment. PBL involves team formation where necessary skills are needed to execute the project. Traditionally, the teams were randomly allocated based on individual preferences. To cab on this issue, preference based model needs few refinements such as skills needs to be identified by the facilitator while the students provide the necessary skill data. This way, students get assigned based on their skill rather than just random allocation. In a worst case scenario for random allocation, a team can end up with a very strong team having high skills or vice versa where a team has all of its members with limited skill or few skills are missing. The group created by skill preference would allow each group to more or less have the same strength and nearly all skills would be present in a group. In this paper, a method is extended from its original to cater for other state-of-the-art optimization techniques rather than just genetic algorithm to find a method that can suit small or large dataset. The objective function takes into account the differences between the total skill set of each group with the average total skill set needed for each group and the missing skill penalty of each group is added. Missing skill penalty is incurred due to not satisfying all the constraints such as non-presence of all the skills in a group. The skill rating allows better selection of members in a software engineering course. The results discussed in this paper are from 5 courses of one university.**

***Index Terms*—Project Based Learning, particle swarm optimization, genetic algorithm, skill based, invasive weed optimization, firefly algorithm**

I. Introduction

Optimization in the field of education is not a new thing, new avenues are sorted to better the teaching and learning process. One such problem involving optimization is Project Based Learning (PBL). PBL is management teaching concept [1] where team dynamics are employed to explore and develop critical thinking. Having groups to work on projects rather than individuals allows cooperation which instills interpersonal skills such as teamwork and collaboration [2]. The success of the project depends entirely on the team dynamics or team composition. Team projects are an important part of any institute where group tasks not only builds teamwork but allows for on time completion of projects [3]. According to Baş and Beyhab [4], project-based learning students have excelled in comparison to traditional learning methods. However, PBL has it share of disadvantages as well, where the wrong formation of a team can be catastrophic.

Team composition is an important part of software engineering where project based learning shift the focus of software engineering from theoretical and conceptual education to practical [5]. It acts as a staple of educational efficacy through projects [5], [6]. To improve student performance, team composition needs to be based on certain principles or criteria. The project or group dynamics allows sharing of ideas, knowledge imparting, and skills strengthening through its strict deadline [3].

Many algorithms are out there which can be used to formulate a solution for this kind of problem but the main issue arises is the team composition and technique to be used. Team composition depends on the pedagogical approach which suits best for the course and the learner. Firstly, which skilled students should be part of a group to maximize its effectiveness, and secondly, optimality. Mahenthiran and Rouse [7] has suggested that the student allocation needs instructor and students sharing the task where instructor assures that the groups have an appropriate mix of skilled students. This allows for better performance of the groups as mentioned in [8]. Eventually, Students themselves prefer allocation to groups based on some skill set [9], [10].

According to [11], [12], diversity in groups nurture group learning where reciprocal teaching is involved [13], [14]. To have similar pace in completing the projects can come about through grouping students by ability [15] and that too reflects creative behavior [16]. According to another research, group performance was better compared to random allocation where students knew each other [7]. Researchers have also found that good students perform better in homogeneous groups, whereas weaker students tend to do better in heterogeneous groups [17], [18]. Learner profile and the context needs to used to form groups where it draws more advantage over random allocation [19]. This type of problem involves optimization where the best possible group needs to be optimized based on certain contexts or learner profile. The contexts are problem depended such that in a software engineering course, one

978-1-7281-2153-6/19/$31.00

of the constraints could be that each group formed needs a specific skilled student such as a programmer to develop the undertaken project.

General and context specific criteria have been used to formulate groups using evolutionary algorithm (EA) in [20] [21] where constraints played a big role. Roland [20]has mentioned in his research that harder and softer constraints can be simulated by having different weights for preferences. Weighing preferences in this type of problem can be simulated through the use of penalty function where penalty gets added to the fitness depending on the lack of weighted preference in individual groups. The penalty function normally serves as a tool to assist in updating of the fitness of the chromosomes in EA where the constraints are either satisfied and/or violated [22]. In the literature, the use of penalty functions has been commonly associated with constrained optimization. One of the advantages of using the penalty functions is its simplicity while the main disadvantage is that penalty reward or violation is usually user defined. Application of the Genetic Algorithm model was done in [23], where good results were sorted on 4 datasets. This gave the insight to use other state-of-the-art techniques in the optimization field.

Particle Swarm Optimization (PSO) is a popular population based optimization technique that operates in a rather constructive matter in resolving optimization problems [24]. Invasive Weed Optimization (IWO) is a meta-heuristic optimization algorithm [25], which represents the ecological behavior of the colonizing weeds in solving different kinds of optimization problems [26]. Firefly algorithm (FA) is rather a new method proposed in [27], is similar to PSO, and loosely motivated by the grouping behavior of fireflies. All the techniques are widely used in different application fields for their stochastic search property. PSO and IWO are not sensitive to the selection of the initial value of the decision variable for problem-solving .

In this paper, we are proposing a method based on FA where the group dynamics needs to be evenly skilled and each student are assigned to a particular group based on certain skills. Preference based system allows optimization algorithms such as PSO, IWO or FA to focus the search for objective vectors towards the region of interest such as the ideal group with all the preference met. While allocating students into their groups, there may be plenty options based on maximally diverse groups, evenly skilled groups, and preference based. Preference can be where friends are part of the same group, distributing subsets students of students, assigning students to specific groups based on skills and other preferences [20].

The rest of the paper is organized as follows. Section II highlights on the formulation of the skilled based group allocation problem. Section III displays the proposed algorithm overview and section IV discusses the experimental setup and results. Section V is on the discussion of the results while section VI concludes the paper with future overviews.

II. Formalization of Skilled Based Group Allocation

In this section, the main components of the formulation will be discussed only. For detail overview of the formulation please refer to [23]. To get a balanced team, the skill set needs to be distributed evenly. One of the important aspects in the formalization is unbiased groups where a group should have all the skill sets present and also share the same strength. Another important aspect is skill rating, which would allow selection of individuals based on their preference. The technique used in this paper is a penalty based model. The skill sets are given a rating based on its importance.

The group allocation problem needs to be minimized based on the penalty based objective function. The objective function is shown in Equation 1.

$$\underset{x\in\Omega}{Min} f(x) \tag{1}$$

where f is the fitness function, x is the input vector and Ω is the parameter space.

The fitness function is provided by problem dependent weight. The weights are similar to the ones used in exam timetabling problem [28]. The rating or preference of particular skill can be calculated by (P_x) as shown in Equation 2.

$$P_x = 2^{L_x - 1} \tag{2}$$

where x=$\{0,1,...,x_{max}\}$ and $L_x$is the preference level.

The weighted skill is shown in Equation 3, which is based on the preference as mentioned above.

$$W_d = 2^{(|L|+1)-P_d} \tag{3}$$

where W_d indicate the weighted preference, $|L|$ is associated levels, P_d is the skill preference. The weight is used to calculate the total skills of individual students and as well as the teams or groups.

In this type of optimization, there is a need for missing skill penalty function. This is used to reward or punish the team. The missing skill penalty (M_j) is shown in 4.

$$M_j = \sum_{x\in\lambda-\lambda_i} W_x \tag{4}$$

subject to

$$\lambda = 1, ... + i : 0 \leq |K|$$
$$||\lambda_i|| \leq ||\lambda||$$

.

where W_x is the weighted preference for a particular skill,$|K|$ is the total number of skills and λ is indicating presence of all the skill set ($\lambda = \{1, 1, 1, 1, 1\}$) while λ_i is subset of skill set present in individual groups (λ_i=$\{1, 0, 1, 0, 1\}$). Therefore, $\lambda - \lambda_i$ will provide the skill set missing from individual groups.

The formulation of the fitness function for discrete dataset is shown in Equation 5, which is the weighted penalty function in this case.

$$\min f(x) = \sum_{i=1}^{G} \left| T_i - \frac{1}{G} \sum_{k=1}^{G} \sum_{x \in \lambda_i} W_x \right| + M_j \quad (5)$$

where x is the input vector, W_x represents the weighted sum of skills, M_j is the missing skill penalty,T_x represents the weighted sum of skills for a group, and G represents the number of groups.

III. Algorithm Overview

The group formulation problem is based on a questionnaire response from students where the responses are either Yes or No (logical in nature). The questionnaire asks for each important skill set required to complete the project where student responses are transformed in binary form to allow it to work with the proposed method.

When each question or asked skill is translated, it is either a 0 (for no) or a 1(for yes). The genome of the chromosome may range from 0000 to 1111 (in binary) which is depended on the number of skills asked in the questionnaire. Each student can then be represented by a bit string, which later forms the dataset.

The proposed algorithm is based on Firefly Algorithm (FA) which is shown in Algorithm 1. This method is compared with one of the popular algorithm namely Particle Swarm Optimization (PSO) and one of the recently developed algorithm Invasive Weed Optimization (IWO) based on group allocation problem where these two algorithms have been briefly explained in 2 and 3 respectively.

The proposed algorithm incorporate preference based optimization where the preferred skill set is made compulsory in terms of weighted objective function as discussed earlier.

Algorithm 1: Preference based approach for student assignment problem using FA

```
Step 1: Initialize population of fireflies
Random partition of n students to associate with
  fireflies G
Step 2: Perform Optimization- Check for preference.
  Each skill needs to be present in each group. Total
  skill of any group should be same. (T_j)
foreach Cycle until maximum iterations do
    foreach Cycle until particles size do
        Evaluate new solutions and update light
          intensity
    end
    Rank fireflies
    Update current best
end
```

In Step 1 of the proposed algorithm, the population is initialized in terms of the fireflies. Each students get randomly partitioned for the first instance in association with fireflies. As for the step 2, the method tries to evaluate the populations where new solutions are created and light intensity is updated.

Algorithm 2: Preference based approach for student assignment problem using PSO

```
Step 1: Initialize population
Random partition of n students into particle positions
  and velocities G
Step 2: Perform Optimization- Check for preference.
  Each skill needs to be present in each group. Total
  skill of any group should be same. (T_j)
foreach Cycle until maximum iterations do
    foreach Cycle until particles size do
        Evaluate fitness function
    end
    Check convergence
    Update particles and positions
end
```

Algorithm 3: Preference based approach for student assignment problem using IWO

```
Step 1: Initialize population of weeds
Random partition of n students into weeds G
Step 2: Reproduction- Check for preference. Each
  skill needs to be present in each group. Total skill of
  any group should be same. (T_j)
foreach Cycle until maximum iterations do
    foreach weed in the population do
        Evaluate fitness function
    end
    Mutation
    Update Global best
end
```

The method tries to find the best individuals through signal system to attract other fireflies and it continues until allocated iterations. Mutation is used to get better solution across the field to benefit the entire population.

IV. Experiments and Results

This section shows the experimental setup and results analysis on best algorithm to use with the proposed formulation of group allocation problem. The experimental setup highlights the parameters used in each of the algorithm and the dataset. Results section showcases the mean, median and best and worst results from the experiment.

A. Experimental Setup

We analyzed the data from four software engineering courses and one project based course at The University of the South Pacific since benchmark datasets were not available in the literature. Four of the courses deal with software development projects, while last dataset looks at project based course requiring knowledge of software development. The skills such as programming, database administration, analysis and design as well as presentation skills is needed to successfully complete and present software-intensive system development projects. Data collection was possible through questionnaire as individual skill capturing was required. Pre-testing and repeated question allowed credibility of the questionnaire.

The dataset used is available online to assist other researchers and can be accessed from [29]. The dataset show individual students skill based on binary input. The skill set indicated for each of the dataset are coder, analysis & design, database and presenter. While android programming is the fifth skill set indicated for datasets 2 and 4. Dataset 1 is comprised of 57 members, dataset 2 has 25, dataset 3 has 31 members, dataset 4 has 20 members while dataset 5 has 234 members. Each algorithm was run 20 times, and the best and worst results (fitness value) together with median and mean of the runs is tabulated. The groups created were based on 3 possibilities; 3 member or 4 or 5 member teams, which are ideal size of group in any project based learning. Table I shows the parameter settings used. The sub-section gives deals about preference and missing skill penalty.

TABLE I
PARAMETER SETTING

Parameter	PSO	IWO	FA
Initial Population size	50	20	20
Maximum iterations	300	300	300
No. of Runs	50	50	50
Inertia Weight	1	-	-
Damping Ratio	0.99	-	0.98
Personal Learning Coefficient	1.5	-	-
Global Learning Coefficient	2.0	-	-
Maximum Number of Seeds	-	5	-
Variance Reduction Exponent	-	2	-
Initial Value of Standard Deviation	-	1	-
Final Value of Standard Deviation	-	0.001	-
Light Absorption Coefficient	-	-	1
Attraction Coefficient Base Value	-	-	2
Mutation Coefficient	-	-	0.2

4.11 Preference and Missing Skill Penalty: The skill rating of any skill is determined by its importance. The most important get first ranking and followed by the rest. As seen in the Table II, the most preferred skill was programmers, therefore, it got preference as 1. The low value indicates that it was most preferred and vice versa for high value. Programmer and android programmer is given preference as 1 while Analysis & Design and Database administrator are given the rating as 2 and presenter skills are given rating 4 as it is the last skill needed to successful showcasing of any software engineering course. The skills are classified under different levels as each level can be associated with multiple skill. The preference plays a big role in this type of optimization problems.

TABLE II
PREFERENCE TABLE

Skill type	Preference	Level
Programmer	1	1
Android Programmer	1	1
Software Analysis and Design	2	2
Database Design & management	2	2
Research & Presentation	4	3

To get a balanced group, the fitness function checks the total skill count of each group, the presence of each skill, and the number of members in a group. Each skill is given a rating based on preferred skill type needed for the group projects and if a preference is not present, there is a missing skill penalty applied to the fitness function as shown in Equation 4. Missing skill penalty is penalty incurred due to lack of a particular skill in a group. The amount of penalty is determined by the type of skill. The most preferred skill has more penalty than a less preferred skill.

B. Results

This subsection reports on analysis of the algorithm. The Tables III and VIII shows the results obtained for the five datasets.

In the Table III, the group ranges from 3 members to 6 members. The mean, median, best and the worst of the 50 runs is shown. The best results were seen FA model. The best result got was less than 2 and worst was less than 10 for 4 member group. PSO had slight better performance than IWO.

In the Table IV, the group ranges from 3 members to 5 members where the number of students in the course was 25. The mean, median, worst and the best of the 50 runs is shown. The results were similar for all the three methods but again FA outperformed other two in terms of 3-4 member group.

The Table V displays results for dataset 3. The group ranges form 3-6 where mean, median, best and worst results for 50 runs is shown. Results were similar for all the methods where FA again did bit better than the other 2. PSO also had better performance than IWO except for 5-6 member. Fitness of the all the methods got better as the number of members increased.

The Table VI displays results for dataset 4. The group ranges form 3-6 where mean, median, best and worst results for 50 runs is shown. The number of members for this particular class was just 20 students. Results were similar for all the methods where IWO performed better than rest of the methods in terms of the mean values of first two instance of the groups. Fitness of the all the methods got better as the number of members increased in groups.

TABLE III
THE PROPOSED METHOD'S PERFORMANCE IN DATASET 1

Method		Members		
		3	4-5	5-6
FA	Best	**36.63**	**1.86**	**3.27**
	Median	**37.63**	**1.86**	**4.91**
	Worst	**38.63**	**5.86**	**8.18**
	Mean	**37.68**	**2.26**	**5.15**
IWO	Best	47.63	9.57	8.55
	Median	59.63	17.29	12.55
	Worst	68.05	27.57	17.82
	Mean	58.57	18.51	12.72
PSO	Best	38.63	3.71	**3.27**
	Median	42.63	7.71	6.55
	Worst	51.05	19.57	10.18
	Mean	43.32	8.38	7.24

TABLE IV
THE PROPOSED METHOD'S PERFORMANCE IN DATASET 2

Method		Members		
		3-4	4-5	5
FA	Best	**54.00**	**34.67**	**25.00**
	Median	**54.00**	**34.67**	**25.00**
	Worst	56.00	**34.67**	**26.00**
	Mean	**54.30**	**34.67**	**25.40**
IWO	Best	**54.00**	**34.67**	**25.00**
	Median	55.00	**34.67**	26.00
	Worst	56.00	**34.67**	**26.00**
	Mean	54.70	**34.67**	25.60
PSO	Best	**54.00**	**34.67**	**25.00**
	Median	55.00	**34.67**	**25.00**
	Worst	**55.00**	**34.67**	**26.00**
	Mean	54.65	**34.67**	25.50

TABLE V
THE PROPOSED METHOD'S PERFORMANCE IN DATASET 3

Method		Members		
		3-4	4-5	5-6
FA	Best	**7.33**	**2.67**	**1.60**
	Median	**8.11**	**2.67**	**1.60**
	Worst	**8.11**	**4.67**	3.60
	Mean	**7.76**	**2.87**	**1.78**
IWO	Best	**7.33**	**2.67**	**1.60**
	Median	**8.11**	4.00	**1.60**
	Worst	9.33	**4.67**	**3.20**
	Mean	8.21	3.62	**1.78**
PSO	Best	**7.33**	**2.67**	**1.60**
	Median	**8.11**	**2.67**	**1.60**
	Worst	8.89	**4.67**	3.60
	Mean	7.84	3.22	**1.78**

TABLE VI
THE PROPOSED METHOD'S PERFORMANCE IN DATASET 4

Method		Members		
		3-4	4	5
FA	Best	**20.33**	**12.00**	**3.00**
	Median	**20.33**	**12.00**	**3.00**
	Worst	**21.00**	**14.00**	**3.00**
	Mean	20.67	12.80	**3.00**
IWO	Best	**20.33**	**12.00**	**3.00**
	Median	**20.33**	**12.00**	**3.00**
	Worst	**21.00**	**14.00**	**3.00**
	Mean	**20.60**	**12.10**	**3.00**
PSO	Best	**20.33**	**12.00**	**3.00**
	Median	21.00	**12.00**	**3.00**
	Worst	**21.00**	**14.00**	**3.00**
	Mean	20.77	12.50	**3.00**

TABLE VII
COMPARISON OF BEST AND MEAN RESULTS OF 4 MEMBER GROUPS WITH GA METHOD

Method		Data sets			
		1	2	3	4
FA	Best	**1.86**	**34.67**	**2.67**	**12.00**
	Mean	**2.26**	**34.67**	**2.87**	12.80
IWO	Best	9.57	**34.67**	**2.67**	**12.00**
	Mean	18.51	**34.67**	3.62	**12.10**
PSO	Best	3.71	**34.67**	**2.67**	**12.00**
	Mean	8.38	**34.67**	3.22	12.50
GA	Best	**1.86**	**34.67**	**2.67**	**12.00**
	Mean	4.82	34.77	3.36	13.36

TABLE VIII
THE PROPOSED METHOD'S PERFORMANCE IN DATASET 5

Method		Members		
		3	4-5	5-6
FA	Best	**97.62**	**40.69**	12.65
	Median	110.62	45.48	**18.39**
	Worst	124.62	**48.48**	**25.22**
	Mean	112.03	45.28	**18.73**
IWO	Best	375.51	223.14	158.26
	Median	406.36	262.21	178.35
	Worst	449.15	297.28	222.09
	Mean	412.47	258.61	182.21
PSO	Best	181.36	88.55	50.87
	Median	217.31	108.21	66.78
	Worst	259.26	134.14	83.61
	Mean	217.75	108.97	67.72
GA	Best	99.62	40.83	**11.74**
	Median	**105.62**	**43.69**	19.39
	Worst	**113.56**	51.48	25.30
	Mean	**106.58**	**45.08**	18.79

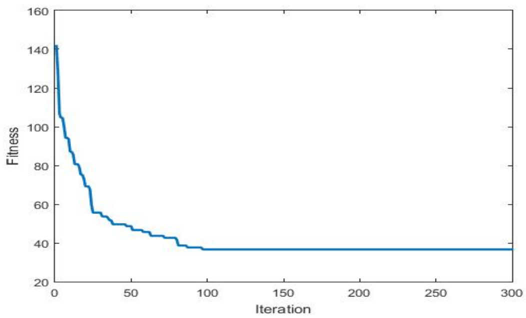

Fig. 1. Typical convergence of fitness through number of iteration on Dataset 1.

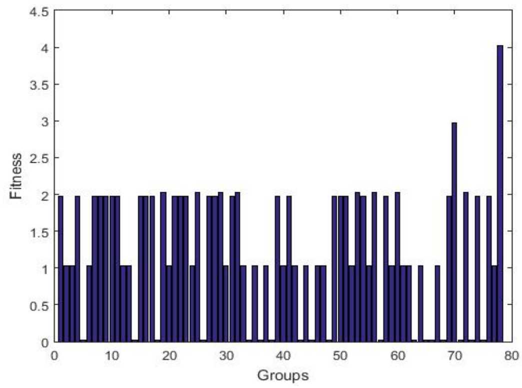

Fig. 2. Typical best fitness of dataset 5

In the Table VII, comparison of best and mean results of 4 member groups is done with previously proposed GA method for group allocation project. Here it was seen that proposed method using FA had better mean values compared to GA, IWO and PSO. The best value for GA and FA were same for all datasets. IWO and PSO had best value same for dataset 2-4 while for dataset 1 it was bit higher in comparison to GA and FA methods. This provided insight to test the algorithm with new dataset which was dataset 5 having 200 plus data.

The Table VIII displays results for Dataset 5 with proposed method in comparison to GA method. The group ranges form 3-6 where mean, median, best and worst results is shown. The results are more towards GA and FA methods. FA gave best values for 3 and 4 member group while GA gave best value for 5 member group. Ga gave low mean values for 3 and 4 member groups while FA gave for 5 member group. The result range was same for GA and FA while PSO had high values and worse was seen for IWO method.

The Figure 1 displays how the solution of a dataset converges with number of iteration. Figure 2 shows the typical fitness values of all the groups formed for a dataset using FA method. Figure 3 shows comparison of convergence and fitness values for FA, IWO and PSO models. Lower fitness, better performance, which is ideal to FA algorithm.

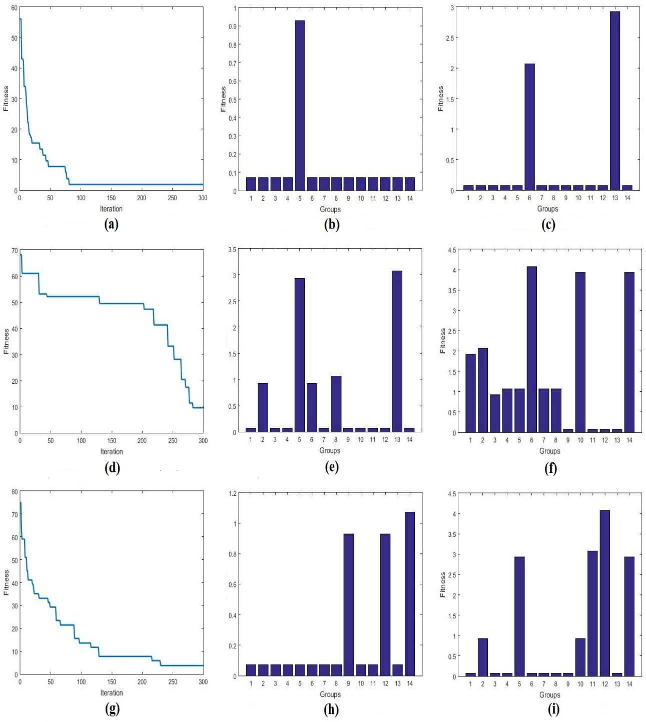

Fig. 3. Typical fitness values with convergence for a Dataset having 4 member group. (a)-(c) shows results for FA, while (d)-(f) shows for IWO and (g)-(i) for PSO model. First is convergence, second is best fitness and third is worst fitness for each method.

V. Discussion

The performance for all three algorithms (FA, PSO and IWO) with group allocation problem will be discussed in this section. It has been seen nearly in all the datasets, Firefly based model performed better in comparison to the other two methods. For all the datasets, the performance of the algorithms converged better as the number of group members was increased. This allowed for the better spread of individuals as for Dataset 2 to 4 had a limited number of students which resulted in a lack of skills. The larger the student number the better chances to get the students evenly arranged if the necessary skills are available. All the groups need to have more or less the same strength in terms of total skills, as well as all skills should be present in a group and most importantly the skills are based on preference. The preferred skill was allocated a higher skill rating.

In dataset 1, it was seen that the FA method outperformed the rest of the methods as it was able to give better performance. Four to five member group had the lowest fitness value of less than 2 indicating that missing skill penalty was few as all the groups had similar strength compared to the other. The results on dataset 2 for all the three methods were similar in most cases, indicating all methods were able to find the best results. The dataset was the smallest set of all the five datasets and it shows smaller dataset was solved well with the all the methods.

In dataset 3, results were pleasing for all the methods but FA method was best one out. where it performed best in all the number of groups from 3 to 5 member. As for the dataset 4, it was seen the IWO method had the best result compared to PSO and FA. Other two methods did well as well, just in two cases the mean value was bit higher.

The results of dataset 1 to 4 were compared with the existing method proposed in [23] where the same dataset was used with GA method. FA method outperformed in all the cases where GA method's mean was a bit higher in all the cases. The performance indicates FA method is a good solution for this type of problems. To further investigate the algorithm, dataset 5 was introduced where the number of students to group was 234. It was seen that FA and GA methods did well in this dataset while PSO and IWO had the worst performance. The structure used by FA and GA works well with this type of problem where penalty function is used penalize missing skill of a group.

For more diversity within groups, the number of skills types and the number of group members is very much important. For example, if there are more students to allocate into groups but the number of important skill is limited, there may be a case where not all groups will get an important skilled individual due to skill limitation. The strength of the of groups may come to same but a coder cannot be replaced with two database administrators.

By looking at Figure 3, further analysis can be done on the performance of individual methods. The rows represent methods, starting with FA and ending with PSO, showing convergence, best fitness and worst fitness of the groups created for Dataset 1 for 4-5 member groups. The best and worst result of FA was similar just that the fitness of one more group was higher in the worst case scenario. The rest of the groups formed had the same fitness values indicating the best solution. Fitness value was not the same throughout the groups due to the fact, a group ended up with 5 members rather than 4 as the result of an odd number of students in the class. If we look at the performance of IWO method which comprises the convergence, best and worst cases, it is shown in the middle of the figure labeled from d to f respectively. The best result is worse than the worst case of FA. The method tries to solve the problem with 300 iteration for this small set of data but due to penalty function usage, the method's performance is not that good. This approach is not appropriate for this type of problem with higher student ratio. Same can be seen for PSO method, even though its performance is better than IWO but its performance is not as good as FA.

If we tend to look at the results again, in some cases the worst case gives totally unbalanced groups which are also possible with random allocation of students in most cases. There can be a possibility, the performance of a group will not be at its peak due to its members or worse case scenario not having those important skilled people such as programmers in a software development course. Getting optimized results is far the best result than manually adjusting the groups of 100 to 1000 students or letting the students form their own groups.

This results are proof of concept and at the preliminary level as more in-depth analysis needs to be done to allow the best algorithm to solve such problems in the nip of time.

The results cannot be compared to existing literature as benchmark dataset availability is not there and more importantly the source code of the existing literature is not available. However, comparison to GA method was possible, which sparked new avenues.

VI. Conclusions

This paper employed the Firefly Algorithm (FA) to solve the Group Assignment Problem (GAP) where comparison was done using Particle Swarm Optimization (PSO) and Invasive Weed Optimization (IWO). These algorithms were used to find the optimal groups for a different set of data based on weighted skills. Experimental results showed that all methods had similar performance on smaller datasets while better performances came from FA method. For a large set of data, again FA performed better when compared to the other two methods. This allowed comparing the results of the dataset with an ideal group with already existing GA method for this type of problem. FA method again showed that it had better performance than GA method in terms of mean value while their best results were same. Further comparison was done using the last dataset having more than 200 students where GA and FA had similar results.

The results obtained are proof of concept and is at preliminary stage since it is based on four software engineering courses and one project based course from our University. The paper can be extended to cater for large sets of data with depth analysis on the use of algorithms. The results will be statistically tested to proof the claims and show robustness of the proposed method in future work.

References

[1] M. L. Moliner, T. Guraya, P. Lopez-Crespo, M. Royo, J. Gamez-Perez, M. Segarra, and L. Cabedo, "Acquisition of transversal skills through pbl: a study of the perceptions of the students and teachers in materials science courses in engineering," *Multidisciplinary Journal for Education, Social and Technological Sciences*, vol. 2, no. 2, pp. 121–138, 2015.

[2] L. Thorley, R. D. Gregory, and R. Gregory, *Using group-based learning in higher education*. Psychology Press, 1994.

[3] Y. Sunaga, H. Washizaki, K. Kakehi, Y. Fukazawa, S. Yamato, and M. Okubo, "Relation between combinations of personal characteristic types and educational effectiveness for a controlled project-based learning course," *IEEE Transactions on Emerging Topics in Computing*, vol. 5, no. 1, pp. 69–76, 2017.

[4] G. Baş and Ö. Beyhab, "Effects of multiple intelligences supported project-based learning on studentsâ achievement levels and attitudes towards english lesson," *International Electronic Journal of Elementary Education*, vol. 2, no. 3, pp. 365–386, 2017.

[5] S. Sonavane, M. Rathi, and S. Patil, "Practice based closure quality loop optimization in teaching learning process: A case study in software engineering design," *International Journal of Education and Learning Systems*, vol. 3, 2018.

[6] L. Springer, M. E. Stanne, and S. S. Donovan, "Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis," *Review of educational research*, vol. 69, no. 1, pp. 21–51, 1999.

[7] S. Mahenthiran and P. J. Rouse, "The impact of group selection on student performance and satisfaction," *International Journal of Educational Management*, vol. 14, no. 6, pp. 255–265, 2000.

[8] Y. Eshel and R. Kohavi, "Perceived classroom control, self-regulated learning strategies, and academic achievement," *Educational Psychology*, vol. 23, no. 3, pp. 249–260, 2003.

[9] K. J. Chapman, M. Meuter, D. Toy, and L. Wright, "Canât we pick our own groups? the influence of group selection method on group dynamics and outcomes," *Journal of Management Education*, vol. 30, no. 4, pp. 557–569, 2006.

[10] P. Kelly, "Achieving desirable group-work outcomes through the group allocation process," *Team Performance Management: An International Journal*, vol. 14, no. 1/2, pp. 22–38, 2008.

[11] S. B. Feichtner and E. A. Davis, "Why some groups fail: A survey of students' experiences with learning groups," *Organizational Behavior Teaching Review*, vol. 9, no. 4, pp. 58–73, 1984.

[12] T. E. Muller, "Assigning students to groups for class projects: An exploratory test of two methods," *Decision Sciences*, vol. 20, no. 3, pp. 623–634, 1989.

[13] A. S. Palinscar and A. L. Brown, "Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities," *Cognition and instruction*, vol. 1, no. 2, pp. 117–175, 1984.

[14] B. Rosenshine and C. Meister, "Reciprocal teaching: A review of the research," *Review of educational research*, vol. 64, no. 4, pp. 479–530, 1994.

[15] J. A. Kulikand and C.-L. C. Kulik, "Effects of ability grouping on student achievement," *Equity and Excellence in Education*, vol. 23, no. 1-2, pp. 22–30, 1987.

[16] R. H. Rutherfoord, "Using personality inventories to form teams for class projects: a case study," in *Proceedings of the 7th conference on Information technology education*. ACM, 2006, pp. 9–14.

[17] N. M. Webb, K. M. Nemer, A. W. Chizhik, and B. Sugrue, "Equity issues in collaborative group assessment: Group composition and performance," *American Educational Research Journal*, vol. 35, no. 4, pp. 607–651, 1998.

[18] M. Lejk, M. Wyvill, and S. Farrow, "Group assessment in systems analysis and design: a comparison of the performance of streamed and mixed-ability groups," *Assessment & Evaluation in Higher Education*, vol. 24, no. 1, pp. 5–14, 1999.

[19] M. Muehlenbrock, "Learning group formation based on learner profile and context," *International Journal on elearning*, vol. 5, no. 1, p. 19, 2006.

[20] R. Hubscher, "Assigning students to groups using general and context-specific criteria," *IEEE Transactions on Learning Technologies*, vol. 3, no. 3, pp. 178–189, 2010.

[21] K. Baker, S. Powell *et al.*, "Methods for assigning students to groups: A study of alternative objective functions," *Journal of the Operational Research Society*, vol. 53, no. 4, pp. 397–404, 2002.

[22] C. A. C. Coello, "Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art," *Computer methods in applied mechanics and engineering*, vol. 191, no. 11, pp. 1245–1287, 2002.

[23] R. Nand, A. Sharma, and K. Reddy, "Skill-based group allocation of students for project-based learning courses using genetic algorithm: Weighted penalty model," in *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, Dec 2018, pp. 394–400.

[24] D. P. Rini, S. M. Shamsuddin, and S. S. Yuhaniz, "Particle swarm optimization: technique, system and challenges," *International journal of computer applications*, vol. 14, no. 1, pp. 19–26, 2011.

[25] H. S. Rad and C. Lucas, "A recommender system based on invasive weed optimization algorithm," in *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*. IEEE, 2007, pp. 4297–4304.

[26] A. R. Mehrabian and C. Lucas, "A novel numerical optimization algorithm inspired from weed colonization," *Ecological informatics*, vol. 1, no. 4, pp. 355–366, 2006.

[27] X.-S. Yang, "Firefly algorithms for multimodal optimization," in *International symposium on stochastic algorithms*. Springer, 2009, pp. 169–178.

[28] R. Qu, E. K. Burke, B. McCollum, L. T. Merlot, and S. Y. Lee, "A survey of search methodologies and automated system development for examination timetabling," *Journal of scheduling*, vol. 12, no. 1, pp. 55–89, 2009.

[29] R. Nand, "Software-engineering-project-questionnaire-dataset," https://github.com/ravneil/Software-Engineering-Project-Questionnaire-Dataset, 2018.