Dynamic Re-Composition of Learning Groups Using PSO-Based Algorithms

Zhilin Zheng
Department of Computer Science
Humboldt University of Berlin
Berlin, Germany

zhilin.zheng@hu-berlin.de

Niels Pinkwart
Department of Computer Science
Humboldt University of Berlin
Berlin, Germany

niels.pinkwart@hu-berlin.de

ABSTRACT

In collaborative learning contexts, the problem of automatically forming effective learning groups gets considerably complex with larger class sizes, e.g. in MOOCs. Additionally, group dynamics caused by high dropout rates currently observable on online open course platforms poses challenges to learning group formation strategies. To address these problems, this paper presents PSO-based algorithms to dynamically re-compose learning groups. In addition to static grouping criteria (such as MBTI personality types), the algorithms take into account factors of the group success rate and group satisfaction during re-composition. We carried out simulations based on randomly generated sample data. The experimental results show that the proposed approach performs better than traditional exhaustive or random methods.

Keywords

Group Formation; Collaborative Learning; Group Composition; Group Dynamics

1. INTRODUCTION

The concept of re-composing learning groups was introduced by Oakley et al. [4]. They dissolved dysfunctional teams to re-form more effective teams. However this is just one single motivation for re-forming learning groups. There is another important reason that should not be forgotten in the light of the growing popularity of massive open online courses (MOOCs): the dropout rate. According to Dung Clow's findings, the dropout rate on MOOC platforms is considerably higher than in traditional education [2]. Only 3% of the initial participants took the final exam in a bioelectricity MOOC Duke offered through Coursera [5]. In collaborative learning contexts, this high dropout rate may cause learning groups to collapse. Therefore, it is crucial to re-compose learning groups in order to enable an effective collaborative learning setting also in later parts of a course when many participants might have left.

The rest of the paper is organized as follows. The following two sections describe our research methods and the proposed approach. We then present the simulation results. Finally, the last section concludes this paper.

2. METHOD

In this paper, we assume that a class S is composed of a given number of n students, $S = \{s_1, s_2, \dots s_n\}$. Before taking a course, instructors must divide these n students into r groups, $G = \{g_1, g_2, \dots g_r\}$. Each student can only be a member of a single group. G is the initial group formation which can, for instance, be formed by diversifying MBTI personality and distributing even gender. Subsequently, these r groups of students are instructed to complete their first group tasks. When they finish, every group's work is rated, $SR = \{SR_1,\ SR_2,\ \dots, SR_r\}$. From the rating data, we then estimate the pair success rate PS_{ij} (i.e. if s_i and s_j are in group g_k, then $PS_{ij} = SR_k$). In parallel and in addition to the performance rating, the participating students are invited to state their personal satisfaction rating, SA_{ij}, with respect to their teammates. SA_{ij} stands for s_i's subjective satisfaction rating about working in one group with s_j. The satisfaction rating indicates how much one student is willing to work with each of his teammates. When the satisfaction rating is low, the student will very likely not want to stay in the same team with his counterpart. Between group tasks, we also assume some certain percentage of students dropping out from the course. Then, in the next group task, we intend to re-compose the remaining students into learning groups aiming at meeting the initial grouping criteria as well as maximizing group success rate and pair satisfaction. We then follow this strategy to re-compose learning groups task by task.

3. PSO-BASED APPROACH

In order to solve our group re-composition problem, a Discrete Particle swarm optimization (DPSO) algorithm is proposed in this study which was previously introduced to the manufacturing cell design problem [3] and the travelling salesman problem [1]. We use a list representation for a group formation: *n* students are simply permutated in a list of length *n*. The PSO starts with initial solutions which are called particles, then updates these initial solutions and searches for the optimal solution iteratively. The velocity vector v_k^{t+1} which is used to update a particle P_k for the next iteration can generally be calculated using (1).

$$\mathrm{v}_k^{t+1} = \mathrm{c}_1 * \mathrm{v}_k^t + \mathrm{c}_2(\mathrm{P}_{k,best} - \mathrm{P}_{k,current}) + \mathrm{c}_3(\mathrm{G}_{best} - \mathrm{P}_{k,current}) \quad (1)$$

In (1), t stands for the number of the current iteration and k indicates the number of the updated particle. $P_{k,current}$, $P_{k,best}$ and G_{best} indicate the current state of P_k, the personal best prior state of P_k and the global best particle state. c_1, c_2, c_3 are learning coefficients. Representation-wise, a velocity vector v_k is a set of pairwise permutations (i, j) that will be used to update P_k^t to P_k^{t+1} as shown in (2).

$$\mathrm{P}_k^{t+1} = \mathrm{P}_k^t + \mathrm{v}_k^{t+1} \quad (2)$$

In DPSO two fitness functions should necessarily be designed to evaluate the quality of each group formation at the initial stage and re-composition phases respectively, as shown in (3) and (4). Here, $\mathrm{D(g_i)}$ is an indicator of diversity of MBTI personality and gender distribution in a learning group (the larger the better).

$$Z_{ini}(P_k) = \frac{\sum_{i=1}^{r} D(g_i)}{r}, 0 \le Z_{ini}(P_k) \le 1 \quad (3)$$

$$Z_{re}(P_k) = \frac{\sum_{i=1}^{r} w_d \times D(g_i) + w_{sr} \times SR(g_i) + w_{sa} \times SA(g_i)}{r} \quad (4)$$

The complete DPSO algorithm is described in Figure 1.

```
------------------------------------------------------------------------------
Step1: initialize a population of N particles  randomly;
Step2: do {
            evaluate the fitness of each particle by the equation (3);
            for each particle
                update the personal best and global best;
                update velocity by the equation (1);
                update the particle by the equation (2);
            end for
            }
       while (the maximal number of iteration is not reached)
Step3: output the global best;
Step4: do{
            Do Step1, Step2 and Step3 but use the equation (4) to
            evaluate the fitness of each particle instead of (3);
            }
       while (all group tasks not finished)
------------------------------------------------------------------------------
```

Figure 1. DPSO algorithm to re-compose groups

4. SIMULATION RESULTS

As shown in the formula (4), the group quality is calculated based on the MBTI and gender diversity, the group success rate and the group satisfaction rate. The impacts of these three factors are controlled by three weights (i.e. w_d, w_{sr}, w_{sa}). Basically, we have two ways to determine the weight factors. One way is to use fixed weights (possibly gained by experience of through systematic research and test).We simply set $w_d = 0.3$, $w_{sr} = 0.3$, $w_{sa} = 0.4$ for our tests. The other way is to define the weight factors w_d and w_{sr} adaptive to the students' co-working experience. If one group of students has worked together for many times, we can emphasize their previous group success rate and pair satisfaction and reduce consideration of their personal traits diversity. Technically, we set $w_d = 0.3 \times (1 - \alpha^{1/3})$, $w_{sr} = 0.3 \times (1 + \alpha^{1/3})$, $w_{sa} = 0.4 \times (1 + \alpha^{1/3})$. α is a co-working experience factor. We conducted an experiment to test the two methods (fixed vs adaptive weights) on a randomly generated dataset in comparison to two traditional methods, the random method and the exhaustive method.

4.1 Synthetic Data

Participating students' personal traits which exactly contain gender and MBTI personality are typically collected via online surveys. In our research, we generated this data randomly (i.e. each student was randomly assigned a gender information and an MBTI personality type). We designed 4 data sets (made up of 150, 300, 900 and 3000 students respectively) and used this dataset for 8 group re-compositions to test our algorithms. At the stages of group re-compositions, we modeled a dropout rate of 40%, 20%, 10%, 8%, 6%, 4%, 2%, 2% from the first group re-composition to the last one. Group performance and pair satisfaction were also randomized.

4.2 Performance Analysis

The proposed DPSO algorithm has been implemented in MATLAB and tested on the synthetic data illustrated in the previous subsection. The group size in the experiment was set to three. The simulation was conducted on a personal computer with an Intel(R) Core(TM) i7-4600U CPU 2.10GHz and 8GB RAM. We evaluate the DPSO algorithm's performance by computational time and quality of grouping (as measured by the fitness value). As a result, the DPSO algorithm can achieve a near-best solution to our re-composition problem in comparison to the exhaustive method, and runs considerably faster (for 3000 students, the time cost on our machine was just 11 minutes at maximum). The fixed-weights method performs closely similar as the adaptive-weights method in terms of group quality and time cost. As anticipated, the adaptive-weights method composes fewer groups with low satisfaction pairs by comparison of the fixed-weights method, as shown in Table 1 (the percentage indicates how many groups contain a pair of members with a pair satisfaction lower than 0.3).

Table 1. Low pair satisfaction percentages

	fixed-weights				adaptive-weights			
	150	**300**	**900**	**3000**	**150**	**300**	**900**	**3000**
Comp.	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1st re-Comp.	50.7%	60.5%	62.7%	68.5%	28.0%	43.3%	51.8%	62.1%
2nd re-Comp.	57.3%	56.8%	63.3%	66.6%	31.0%	44.0%	49.3%	63.0%
3rd re-Comp.	42.0%	54.3%	56.4%	65.5%	28.0%	30.0%	47.8%	60.0%
4th re-Comp.	39.1%	48.6%	57.3%	63.8%	23.6%	30.9%	44.8%	56.3%
5th re-Comp.	25.0%	48.1%	52.1%	62.7%	10.0%	25.0%	37.9%	54.5%
6th re-Comp.	45.0%	44.2%	50.0%	65.6%	6.7%	18.3%	37.8%	50.7%
7th re-Comp.	52.0%	45.0%	51.7%	60.6%	40.0%	8.0%	37.3%	49.2%
8th re-Comp.	47.5%	37.5%	47.1%	57.9%	25.0%	12.5%	26.7%	46.5%

5. CONCLUSION AND FUTURE WORK

In this paper, we presented a new method for dynamically re-composing students into learning groups by taking into account both (static) personal characteristics, dynamic data (student group success and satisfaction) and student dropout rates. We also proposed a DPSO algorithm to dynamically re-compose collaborative learning groups based on the method. The proposed algorithm is able to search for the near-best solution to our group re-composition problem in an acceptable computational time as compared to the exhaustive method. Additionally, the adaptive-weights method is able to largely reduce the violation of low pair satisfaction. In our future research, we will test this algorithm against real data collected from large online courses.

6. ACKNOWLEDGMENTS

This research is supported by the China Scholarship Council (CSC).

7. REFERENCES

[1] CHANGSHENG, Z., JIGUI, S., YAN, W., and QINGYUN, Y., 2007. An Improved Discrete Particle Swarm Optimization Algorithm for TSP. In *IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology Workshops, 2007*, 35-38.

[2] CLOW, D., 2013. MOOCs and the funnel of participation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (Leuven, Belgium2013), ACM, 2460332, 185-189.

[3] DURAN, O., RODRIGUEZ, N., and CONSALTER, L.A., 2008. A PSO-Based Clustering Algorithm for Manufacturing Cell Design. In *First International Workshop on Knowledge Discovery and Data Mining, 2008. WKDD 2008.*, 72-75.

[4] OAKLEY, B., FELDER, R.M., BRENT, R., and ELHAJJ, I., 2004. Turning student groups into effective teams. *Journal of student centered learning 2*, 1, 9-34.

[5] RIVARD, R., 2013. Measuring the MOOC dropout rate. *Inside Higher Ed, March 8, 2013*.