A New Algorithm for Balancing Group Population in Collaborative Leaning Sessions

Aiman Turani

Faculty of Computer Science and Engineering
Taibah University
Media, KSA

Jawad Alkhatib

Faculty of Computer Science and Engineering
Taibah University
Media, KSA

Abstract—**Proper group formation is essential in conducting a productive collaborative learning session. It specifies the internal structure that the collaborating groups should have based on roles. The Group formation is a dynamic and a challenging component. In some cases, more than one group formation is made during a single session and more than one role a single participant undertakes. Group population is another essential process that follows group formation. Group population is concerned with the process of assigning participants to groups and roles. There are several group population methods that are used in Computer Support for Collaborative Learning environment.**

The main challenge in group population is the partial filling situation. Partial filling is happened when some participants fail to attend their assigned groups at the start of a session. Partial filling could be caused by various reasons. It could be caused by human's mistakes or by technical faults. In this paper, a correction algorithm is described to balance back the groups' participation levels. This algorithm is based on three main phases: Group Elimination, External Transfer, and Internal Transfer phase. The algorithms of these phases are fully described in this paper.

Keywords—Group formation; Group Population; CSCL; Collaboration Techniques

I. INTRODUCTION

Usually, the initial step at the start of any collaborative learning session is to allocate participants into groups and to explain them the nature of roles that they need to do [1]. A session could be composed from a single group or from multi-groups. In the corporate and business environment, a specialized team's members are usually allocated in a single group. In the educational environment, large learning sessions such as workshops, seminars, courses, etc., usually distribute their participants into multi-groups to perform group-based activities. Group collaboration is considered an effective way to transfer valuable knowledge among learners [2]. There are many collaboration techniques, such as Brainstorming, Debate, etc. that are widely used, which are based on defining certain roles and groups' settings [3].

Assigning participants to their groups and roles is a challenging task. Instructor needs to decide how and when roles and group should be populated with participants. Some methods are based on participants themselves to choose, or on instructors or even systemically where participants are chosen randomly or selectively based on their characteristics such as age, sex, level of education, expertise, etc.

It is even a more challenging task when forming these groups in a virtual environment [4]. In face-to-face it is easier for an instructor to direct participants during groups' construction and to engage directly in fixing any grouping issue if it occurs.

II. GROUP POPULATION

Not many Computer Support for Collaborative Learning (CSCL) tools specify how participants will be actually populated in their groups. Some tools allow instructors to create groups based on their students' profile [5]. Other tools assign participants randomly to groups on the base of first come first assigned. Usually, these group formation tools do not take in consideration roles found within collaboration sessions [6]. In addition, they tend to limit the selection flexibility by relying on certain algorithms, such as genetic algorithm [7], and excluding students and instructors from making their own selection.

To clarify the complexity nature of group population, three key issues are listed: population assigner, population time, and population mode. Firstly, the assigning method should be specified early by the owner of the session. In general, there are three common approaches in assigning participants to groups. The assignment could be done by either the participant himself, by the facilitator, or by the system. In the first approach, a participant will simply choose his role and group [8]. In the second approach, the facilitator or the instructor will assign participants to groups. In the last approach, there are two means for the system that it could take, randomly based on first come first assigned, or selectively based participants' profiles data (age, sex, education level, specialties, friendship, geographical background, etc.) [9].

In addition to the assignment method, there are also different population timings. Population could happen just before the session starts or sometime before the session starts. In some cases, the instructor may need longer time to manually allocate participants in their groups especially at large classes. In other cases, he wants participants to know their roles and groups in advance to prepare themselves, or to be able to accommodate any request of changing roles or groups [10].

Finally, there are two identified population modes which are sequential and parallel. In the sequential mode, the system would open up groups for population one by one. In the

parallel mode, all groups would be opened at once. Many CSCL systems use the sequential population at the start of a session allowing arriving participants to join their session in a linear manner. The parallel mode in the other hand provides more selection's flexibility. For example, instructor could start filling groups by first assigning the most critical role at all groups to key-participants and then turns to another role. In another example, a participant would like to select a certain group to join his friends for instance.

III. POPULATION RISKS

There are several populating risks that lead to the groups' partially-filled situation. The partially-filled situation usually causes undesired events to occur. Examples of these events could be the inability to start a session due a large number of absentees, a certain group cannot function properly due to a critical role that is missing or that suddenly leaves, unbalanced group's outputs due to some groups have maximum attendants while other suffers from shortages in attendants, etc.

The focus in this section is on the reasons that might cause partial groups' filling. Firstly, it starts with the last-group partial filling scenario. Even in case of a complete attendance, there is a still high chance that the number of attending participants will not match exactly the default groups' sizes. In other word, the division result will not be always an integer causing the last group to be in short. Another common reason is when some participants fail to join their groups. This can be caused by various reasons such as, carelessness, forgetting, emergencies, technical fault, etc. These undesired incidents could leave some groups with some vacancies or even in extreme cases a scattered population throughout all groups. Late comers issue also needs to be addressed. Would the system allow them to join and if they join, which group and role they could have, and after what point in time they should be banned.

IV. FLEXIBLE GROUP STRUCTURE

There is a need to design a flexible group structure that could tolerate population faults. A group structure within collaborative learning environments depends heavily on roles. What types of roles that are needed to carry out collaborative activities, how many participants should be assigned to a specific role, etc.? These questions should be addressed by an instructor during group's formation phase [11]. To enable flexibility, a further question should be asked, that is which roles are critical and which roles are flexible. To answer this question, the group structures of common collaboration techniques should be analyzed. A certain role is considered a flexible role if its size could vary otherwise it is a critical role. So in order to enable role's flexibility, a maximum and a minimum size should be also defined. For instance, instead of defining a rigid structure for a Brainstorming technique that may include 1 chairperson and 4 participants, it could have this flexible setting: (1 chairperson, 2-7 (5) participant). In this setting the participant role default size is 5 but this size could range from minimum 2 to maximum 7. A simple flexibility ratio could be obtained by using this formula:

Fr = (max group size - min group size)/ default group size

For the above Brainstorming technique the Fr would equal 0.8 (7-2 /6). The following table includes the flexibility ratios for some common collaborative techniques with suggested roles and sizes.

TABLE I. FLEXIBILITY RATIO FOR SOME COMMON COLLABORATIVE TECHNIQUES

Collaboration Technique	Role1	Role2	Role3	Role4	Flexi ratio
Brainstorming Group Nomination	Chairperson 1	Participant 2 -7 (5)			0.83
Group Discussion Buzz group Round Table Discussion Case study Jigsaw	Participant 2 -7 (5)				1.0
Debate	Prosper 1	Opposer 1	Chairper 1	Audience 2 -7 (5)	0.62
Pyramid Team pair-solo Think-pair-share Panel	Participant 4 -14 (10)				1.0
Pro/Contra	Side1 2 -14 (7)	Side 2 2 -14 (7)			1.7

As shown in this table, most techniques have good flexibility ratios allowing for an appropriate opportunity for group populating correction.

V. POPULATING CORRECTION

A populating correction is needed if one the following cases occur:

- If the attending participants as a total number is not sufficient to fill all groups according to their minimum sizes. --- Group Elimination

- If there is an unbalanced group attendance (highest attendance group - lowest attendance group is more than one). --- External transfer

- If any critical role that is not fulfilled. --- Internal transfer

- If there is an unbalanced roles attendance at a certain group (the difference between [a flexible role attendants – its min size] and [any other flexible role attendants – its min size] is more than one). --- Internal transfer

In this paper, the general correction algorithm that depends on the flexibility within the groups' structure was proposed. There are three correction phases in this process which are: the Elimination phase, the External Correction phase, and the Internal Correction phase. The Elimination phase is started by checking if there are no enough attendants to cover all groups' minimum requirements. In this case, the lowest group in attendance is eliminated and its member are transferred to another group. The External Transfer phase is performed to balance participants' distribution. Groups with extra attendants will donate to groups with lower attendants. The last correction phase is the Internal Transfer phase where attendants of a flexible role are used to replace unattended critical roles or to balance other flexible roles if required.

The following correction algorithm would be carried out just before the start of a collaboration session.

1) Groups will be sorted according to their attendance sizes.

2) The system will calculate the Participation index.

3) If the index has a negative value, then the lowest group will be eliminated and its members will be assigned to the next lowest group.

4) Step 3 and 4 will be repeated until the index value becomes zero or more.

5) To allow a balanced participants' distribution, the system will sort groups again and check if the difference between the lowest group and the highest group is equal to one.

6) If the difference is greater than one, then the highest group will donate one member form its flexible roles to the lowest group.

7) Steps 6 and 7 will be repeated until difference between the lowest group and the highest group is only one.

8) For internal group correction, the system will check if all critical roles are populated and also if there is a balanced distribution between its flexible roles.

The Participation index is computed according the difference between the total number of attendants and the total groups' minimum hosting capability. The following formula is used to compute the Participation index

$$Pi = \sum_{i=1}^{All\ Group} Group\ Att. - Min\ G\ Size$$

For instance, suppose you have 31 participants that are joining Group Nomination technique out of 40 learners. They joined in 6 groups as shown in the first column at the following table.

TABLE II.   GROUP NOMINATION ATTENDANCE EXAMPLE

	Sort	Elimi	Sort	Ex.T	Sort	Ex.T	Sort
5	7	7	**8**	7	**7**	6	7
7	6	6	7	7	7	7	6
4	5	5	6	6	6	6	6
6	5	5	5	5	6	6	6
5	**4**	8	**5**	6	**5**	6	6
4	4 *						Finish

According to the Group Nomination setting, the default group size is 6. The Pi = 31- 6*6 = -5. Since this is a negative index, the lowest group's participants will be eliminated. As shown in the second column, a descending sorting is done then and the lowest group which contains 4 participants is eliminated and donates its participants to the next lowest group as shown in the third column. The index will be checked again, which becomes positive this time (31-6*5). A new sorting will follow and since the size of the highest (8) – lowest(5) > 1, an external transfer from max group to min group will happen as shown in columns 4 and 5. Still, since the difference between the highest and lowest is more than one, a new round of external transfer will happen.

To clarify internal transfer process, suppose that there is a collaboration session that is based on Pro/Contra technique with this group formation structure: (1 chairperson, 1-5(2)

side1, 1-5(2) side2) and one of the groups has this population: (3 side1, 6 side2). Since the critical role is missing, the critical role will first be populated from the flexible role side2 (1 chairperson, 3 side1, 5 side2). The absolute value of (3-2) – (5-2) > 1 therefore another internal transfer is needed for balancing roles' distribution. The new population arrangement would look like this (1 chairperson, 4 side1, 4 side2). No further correction is needed since the critical role is fulfilled and the difference between the flexible roles is not more than one.

VI.   CONCLUSION

Proper group formation and population is essential in conducting collaborative learning sessions. Random population can lead to unbalanced grouping and is unlikely to produce effective groups. Not many CSCL tools support group formation and even less supports group population. In addition, collaborative learning sessions usually have their groups' structure based on roles. Many populating tools do not consider roles and also assume that all participants will attend.

Many population risks were discussed in this paper which causes serious difficulty in conducting a proper collaboration session. These risks cause groups to suffer or even to do not function as required (for instance critical roles are missing). To overcome these risks a new populating correction algorithm was introduced in this paper. The correction algorithm has three main phases. The Elimination phase, External Transfer phase, and Internal Transfer phase. The Elimination phase was used to eliminate some groups if there are no enough attendants. An External Transfer phase was used to provide proper group balancing if needed, where groups with extra attendants would transfer some of its members to groups with less attendants. The Internal Transfer phase was used to balance roles' attendance and to fill unattended critical roles within a particular group. A list of eight steps was introduced in this paper to describe the full population correction algorithm.

This correction algorithm did not specify how the system should handle late comers. In the future, a more comprehensive algorithm will include solving late comers' issue. That correction algorithm will not only keep balancing groups after late comers are arriving but also specifies associated policies within, such as, until what point in time or participation percentage the system should allow late comers to join. Will they construct new groups or distribute them in existing groups using non-critical roles, etc.

REFERENCES

[1] W. Martin, P. Hans-Rüdiger, "Group formation in computer-supported collaborative learning," ACM., pp. 24-31, 2001 [In: Proceedings of the 2001 international ACM SIGGROUP conference on supporting group work].

[2] T. Tobin, R. Daniel, Re-educating the corporation: foundations for the learning organization, John Wiley & Sons, 1994.

[3] G. Peter, "Educational design and networked learning: Patterns, pattern languages and design practice," Australasian Journal of Educational Technology, vol. 21, 2005, pp.82-101.

[4] M. Townsen, M. Samuel, and R. Anthony, "Virtual teams: Technology and the workplace of the future," The Academy of Management Executive, 3th ed, vol. 12, 1998, pp. 17-29.

[5] C. Marreiro, I. Seiji, "Group Formation Algorithms in Collaborative Learning Contexts: A Systematic Mapping of the Literature Wilmax," Collaboration and Technology Lecture Notes in Computer Science, vol 8658, 2014, pp. 199-214.

[6] J. Strijbo, L. Rob, W. Martens, M. Wim, and B. Nick, "The effect of functional roles on group efficiency using multilevel modeling and content analysis to investigate computer-supported collaboration in small groups," Small Group Research 2nd ed, vol.35, 2004, pp. 195-229.

[7] R. Tyson, R. Henry, "Forming Productive Student Groups Using a Massively Parallel Brute-Force Algorithm," WCECS 2013, vol. 1 pp. 23-25, October 2013 [Proceedings of the World Congress on Engineering and Computer Science USA, 2013.

[8] Andersena, D., and G. Richardsona. "Scripts for group model building." *System Dynamics Review* 13, no. 2,1997, pp. 107-129.

[9] Chow, Bob Eng Po. "Regulation of motivation in undergraduate business students learning with the case method: examining an underemphasized aspect of self-regulated learning." PhD diss., Education: Faculty of Education, 2011.

[10] Mezirow, Jack. "Last Gamble on Education: Dynamics of Adult Basic Education." (1975).

[11] P. Dillenbourg, "Over-scripting CSCL: The risks of blending collaborative learning with instructional design," Three worlds of CSCL. Can we support CSCL? ,2002, pp. 61-91.