A Metrics Framework for Evaluating Group Formation

Asma Ounnas
University of Southampton
Building 32, Level 3, Room 3069
SO17 2BJ
+44 (0)23 8059 7208

ao05r@ecs.soton.ac.uk

David E Millard
University of Southampton
Building 32, Level 3, Room 3027
SO17 2BJ
+44 (0)23 8059 5567

dem@ecs.soton.ac.uk

Hugh C Davis
University of Southampton
Building 32, Level 3, Room 3027
SO17 2BJ
+44 (0)23 8059 3669

hcd@ecs.soton.ac.uk

ABSTRACT

Many approaches to learning and teaching rely upon students working in groups. So far, many Computer-Supported Group Formation systems have been designed to facilitate the formation of optimal groups in learning. However, evaluating the quality of automated group formation is not always well reported. In this paper we propose a metrics framework for evaluating group formation based upon a model for constraint satisfaction-based group formation.

Categories and Subject Descriptors

K.3.1 [**Computers and Education**]: Computer Uses in Education – *collaborative learning, computer-assisted instruction, distance learning.*

General Terms

Measurement, Performance, Human Factors, Theory.

Keywords

Efficiency, Group Formation, Optimization.

1. INTRODUCTION

Collaboration has long been considered an effective approach to learning. Research in many disciplines has shown that learning within groups improves the students' learning experience by enabling peers to learn from each other. For a collaborative activity to achieve its learning goal, student have to be allocated to appropriate groups that maximizes their individual learning goals in addition to their groups' goals. In this context, the collaboration goal is usually associated with a set of requirements that have to be satisfied to ensure that the formed groups achieve this goal.

In education, teachers often have to deal with group formation (GF) manually which can sometimes turn into a very complex task especially if the number of students in the class is large. This has led researchers to investigate several techniques for automating this process through the use of computer-supported group formation (CSGF). However, in most existing research, the applications developed are only evaluated against few metrics that do not always reflect their efficiency in forming appropriate

groupings; but rather assume that a positive group output can be interpreted as a success of the followed group formation approach.

We believe that to conduct a group formation efficiency study, different measurement variables are required. In this paper, we discuss the different metrics that have been considered in existing applications. We propose a framework where we describe the set of possible metrics for evaluating group formation efficiency, how to measure them, and the relationship between them. The proposed framework is to be used to analyze the efficiency measurements of a constrained group formation system that we previously introduced in [4].

2. GROUP FORMATION (GF)

There are different approaches to forming groups in education. In [4], we explained that although self selecting formation, where the students get to choose their collaborators, is an effective approach in building networks and communities of students, instructor based formation, where the teacher is the initiator of the formation, is more effective in learning. In this context, the instructor has to form balanced groups of students in terms of expected performance, such that no group will have all the top students, while another have weak ones. In other terms, all groups will have an equal chance to perform well and achieve the goals of the collaborative activity, although this may conflict with the best interest of individual students. Therefore, to form the groups, the instructor has to think about modeling the collaboration goal in a way that satisfies both the task of the collaboration that the students have to achieve as a group, in addition to the individual needs of the students.

In CSGF applications, most research is based on the mathematical modeling of agents' coalition, team, club, or networks formation algorithms. Efficiency in this context is usually measured in terms of the algorithm used for the formation. In this paper, we analyse group formation efficiency in learning regardless of the algorithm used to generate it. This allows the possibility of evaluating different algorithms for the same collaborative goal.

3. RELATED WORK

Since the introduction of CSGF applications, efficiency of group formation systems has been measured in different ways. Examples of these measurements are given in the following literature.

In [6], Soh et al. introduced a multiagent intelligent system called I-MINDS where the instructor, each student, and each group is represented by an agent. The student agents form coalitions dynamically in real-time where each bids to join its favorite group based on their previous performance in group work. To evaluate their application, the authors measured the *effectiveness* of IMINDS in terms of how effective did the instructor and the

students found the system easy to use. The group formation itself was evaluated against the *performance* of the teams, which was measured based on the teams' outcomes (*rewards*) and students' responses in a series of questionnaires that evaluates *team-based efficacy*, *peer rating,* and *individual evaluation*. The first metric was used to determine the perceived success of a team by each member, the second metric was used to help evaluate individuals and how well a team has worked together, while the last metric was used as a self rating measure to check for responses consistency with the peer rating. The agent computes a weighed score to each efficacy question, and each member's score is tallied and multiplied with the peer-based weight.

In [8], the authors introduced a web based system that forms the students groups using knowledge about the collaboration context in real-time. Although the authors did not present any results of evaluating their system, they mentioned that the *comprehensibility of the group formation algorithms* and *the satisfaction of learning groups* to be a key factor of the overall approach acceptance.

Redmond [5] introduces a computer program to aid the assignment of students' projects groups using an instructor-based approach. The students are grouped, using a greedy algorithm, based on the time slot they prefer to collaborate in, and then allocate the projects to the groups based on the members' preferences in the group. The groups are then manually checked for even distribution of grades, and the students who are left unassigned are manually allocated to groups. To measure the efficiency of the formation program, the author introduced an evaluation formula that calculated the *rating of group assignments* by subtracting an *unassigned penalty* representing the program failure in assigned some students from the *sum of all formed group overall rating*. Where a group rating is the product of *rating for the satisfaction of each condition*. The rating of each constraint is calculated separately (context dependent). For example, group project rating is the product of squares of students' rating of the assigned projects.

Another way to evaluating formation efficiency in CSGF systems is to compare the results (the formed groups) generated from the system to manual generated results of the same participants' sample [3]. This technique is usually used in expert recommender systems where the formation is based on finding an expert to help a weaker learner [7]

In a different domain such as industrial organizations, *group effectiveness* is defined as the group's productivity in relation to the needs of the organization [1]. Effectiveness in this context is measured in terms of the group' synergy, performance objectives, skills, use of resources, and innovation. These variables are measured using questionnaires designed to combine the measurements of internal dynamics and external group outputs that facilitate the group's self-assessment [1]. The team performance is assessed in terms of the *mean* and *standard deviation* of individual team members' responses across the six domains. In the same domain, group formation (*team composition*) is calculated in terms of *diversity of membership* [2] as it has been shown that team composition has an influence on the team cohesion, communication, conflicts, and creativity in terms of the team's degree of diversity or homogeneity.

In addition to these metrics, CSGF system can also be evaluated in terms of the traditional system quality measurements such as reliability, and robustness. Other metrics such as formation complexity in terms of *complexity of the algorithm* used to form the groups (i.e. complexity measures of memory space and time) can be considered if the algorithm is defined. In dynamic groups such as coalition formation and social networks formation, one of metrics for evaluating efficiency is the *stability* of the groups, a measurement of the changing membership of the group.

From the literature, we observe that the limitation of most GF applications is the exclusive reliance on the groups' performance measures indicators such as members' responses to questionnaires or post-tests to draw inference about the group formation system performance. From a learning viewpoint at least, group formation efficiency is clearly a multi-dimensional concept, which implies that multiple efficiency indicators besides perceived performance need to be employed.

While different formation constraints might result in different formulas for calculating efficiency, these constraints can be related to GF efficiency in a more abstract way. If so, consideration of defining this relation together with other group formation related measures is required.

4. METRICS FOR GROUP FORMATION
In this study, we are concerned with the evaluation of the group formation in terms of how well the groups were formed rather than how well the groups performed. If the instructor models the collaboration goals for the individuals and the groups as a set of requirements (constraints), then, the success of group formation in this context is defined by the satisfaction of the constraints that define these goals. To facilitate the evaluation of group formation, we propose an analytical metrics' framework that defines what we mean by formation success. To achieve this, we first make the following assumptions: each participant in the class should belong to exactly one group (i.e. non-overlapping group formation), all groups should have the same optimal number of participants (i.e. all groups have a similar size), and all formed groups are stable.

4.1 Definitions
Constraints: we define a constraint as any parameter, variable, or condition that affects the process of the group formation (i.e. in CSGF, the variables that influence the system's decision of allocating participants to appropriate groups). We define the finite set of all possible constraints as $C = \{c_1, c_2, c_3 \dots c_P\}$.

Collaboration Task: we define task *t* as the task of the collaboration activity that the instructor intends for the students' groups to perform. In education, the instructor usually selects a set of collaboration goals $\{\alpha_1, \alpha_2 \dots \alpha_K\}$ that assist in achieving the task (i.e. helps the collaborative activity to achieve maximum learning gain for the groups and individuals participants). For example if the task is a software engineering group project, then example goals can be that all groups are to be balanced in terms of students experience in the field; no female student can be allocated alone in an all-male group; and groups should be multicultural in terms of students' nationalities.

Collaboration Goal: we define a collaboration goal α to achieve task *t* as a set of constraints $\alpha = \{(c_1, v_1), (c_2, v_2) \dots (c_L, v_L)\}$ that the instructor chooses to model the requirements for achieving the goal, where each constraint $c_j \in \alpha$ is associated with a value $v_j \in \mathbb{R}$ that represents the importance of the constraint c_j in achieving the goal α. We define $A \subseteq (C \times \mathbb{R})$ as the finite set of all possible goals. In the example above, the constraints for modeling the goals can be respectively: for each group {average percentage of members' experience average $\simeq$ percentage of

members' experience in the next group}; {n° of females ≠ 1}; {n° of international students ≥ 1, n° of international form the same country < n° of participants in the group}. The last goal is presented with two constraints. The constraints can overlap between the goals with different values v for each goal: $(c_1, v_1) \in \alpha_1 \land (c_1, \check{v}_1) \in \alpha_2$).

Participants: we define the finite set of all individual participants (all students in the class) P = $\{p_1, p_2, p_3 \dots p_M\}$, where M = |P| > 1 is the size of the class.

Groups: we define a group g as a set of participants that have at least 2 elements in it (i.e. $|g| > 1$), where each participant $p_i \in g$ is a *member* of the group. We define the set of all possible groups $G = \{g_1, g_2 \dots g_O\}$ such that $\forall i \in M: G = \mathcal{P}(P) - (\{p_i\} \cup \emptyset)$.

Cohort: we define a cohort as the set of groups $\{g_1, g_2, g_3 \dots g_N\}$ of all participants in the class. We define the set of cohorts G_X where X is the optimal size of the groups to be formed, such that G_X is a pairwise disjoint subset of G ($G_X \subseteq G$) that has cardinality N = M/X and for each element g_i in the set $|g_i| = \pm X$, and X > 1.

Formation: We define a relation R from $\mathcal{P}(A)$ to G_X that maps a set of goals to a set of N disjoint groups. This relation can be any algorithm applied to the set of goals. Therefore, for each set of goals, there is more than one possible set of grouping (allocating students to groups) and therefore more than one possible cohort. This is because although if participants p_i, p_j have similar characteristics in relation to the constraints modeling the goals, then the cohort with p_i in group g_k is not the same as the cohort with p_j in group g_k. We refer to each single grouping of R as a *formation*. We say that a formation is defined by the set of goals that determines the cohort: $form(\alpha_1, \alpha_2 \dots \alpha_K) = \{g_1, g_2, g_3 \dots g_N\}$. Figure 1 shows a simplified diagram of the relationships between the collaboration goals and the group formation. (the values of constraints v are not shown due to space limit).

Productivity: we define productivity as the output of the group or the cohort in relation to the task t that is measured on an absolute scale by the instructor or an examiner of the collaborative activity.
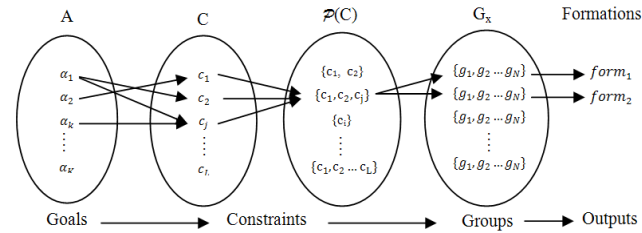

Figure 1 Representation of group formation

4.2 Formation Metrics

4.2.1. Constraint satisfaction Quality
We refer by constraint satisfaction quality to how well the constraints of a goal α_k were satisfied in the formation of the groups (allocation of students). We use this metric to evaluate the *formation quality* later on.

• *Group Constraint Satisfaction Quality*: we use this metric to refer to how well a group g_i is formed in relation to how well the students' allocation (to that group) satisfied a constraint c_j. For each group g_i in the formed cohort, and for each c_j in the set of constraint of α_k ($c_j \in \alpha_k$) we define a function $f_{cg}(g_i, c_j)$ that

determines whether g_i satisfies the constraint c_j such that:
$$f_{cg}(g_i, c_j) = \begin{cases} v_j & \text{if } c_j \text{ is satisfied} \\ 0 & \text{if } c_j \text{ is not satisfied} \end{cases}$$

• *Cohort Constraint Satisfaction Quality:* we use this metric to refer to how well were all the groups formed in terms of satisfying the constraint c_j of goal α_k. We define Cohort Constraint Satisfaction as a function f_{cG} that calculates the degree to which the formed groups are balances (i.e. clustered together) in terms of c_j. Hence we use standard deviation σ to calculate the dispersion of the groups from the mean constraint satisfaction. We define the constraint satisfaction quality to be $f_{cG}(c_j) = \frac{1}{1+\sigma}$ where:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(f_{cg}(g_i, c_j) - \overline{f_{cg}}\right)^2} \text{ and the mean } \overline{f_{cg}} = \frac{1}{N}\sum_{i=1}^{N} f_{cg}(g_i, c_j)$$

Therefore $f_G = 1$ for σ = 0, which is the maximum quality.

4.2.2. Perceived Formation Satisfaction
We use this metric to refer to how well the formation was perceived in terms of participants' satisfaction with the allocations to groups:

• *Individual Perceived Formation Satisfaction*: we use this metric to refer to how pleased is the individual with being allocated a member of the group. Individual satisfaction is usually evaluated using self-assessment questionnaires. Since the questionnaires are usually composed of statements on the Likert scale or the 6 points scale, the satisfactions can be given a weight s_i for each individual p_i where s_i can be the mean of the questions' results.
• *Group Formation Satisfaction:* we use this metric to refer to the individual satisfactions of all the members of the group S_g:

$$S_g = \frac{1}{1+\sigma_{gs}} \text{ where } \sigma_{gs} = \sqrt{\frac{1}{X}\sum_{i=1}^{X}\left(s_i - (\frac{1}{X}\sum_{k=1}^{X} s_k)\right)^2}$$

This metric can be also used to monitor the interactions values of the collaboration such as assistance and conflicts.

• *Cohort Perceived Formation Satisfaction:* Similar to previous analysis, the cohort perceived satisfaction $S_G = \frac{1}{1+\sigma_s}$ where σ_s is the standard deviation of all the groups' satisfactions S_g

4.3 Productivity Metrics

4.3.1. Group Productivity Quality
We refer by quality $Q(t)$ to how well did the group achieve the collaborative task t specified by the instructor. This is a measure of the quality of the group's outcome (sometimes referred to as output or reward) against an absolute scale defined by the instructor or an examiner of the groups' output. In learning, this is usually given in the form of grades or credit to the group. If both the collaboration goal and quality measure are defined by the instructor, then this is a consistent measure.

4.4 Goal Satisfaction Metrics

4.4.1. Goal satisfaction Quality
We use this metric to refer to how well the groups were formed in terms of satisfying a goal α_k within the collaboration task t.
• *Group Goal Satisfaction Quality:* we use this metric to refer to how well a group g_i is formed, in terms of how well the students' allocation (to that group) satisfied the goal α_k. We define Group Goal Satisfaction Quality for goal α_k as a function $f_g(g_i, \alpha_k)$ that calculates the quality of a group g_i in terms of α_k and therefore all constraints of goal $\alpha_k = \{c_1, c_2 \dots c_L\}$ such that

$$f_g(g_i, \alpha_k) = f_g(g_i, c_1, c_2 \ldots c_L) = \frac{1}{L}\sum_{j=1}^{L} f_{cg}(g_i, c_j)$$

• *Cohort Goal Satisfaction Quality:* we refer by Cohort Goal Satisfaction to how well were all the groups formed in terms of satisfying the collaboration goal α_k and hence the constraints that model it. We define Cohort Goal Satisfaction as a function f_G that calculates the degree to which the formed groups are balances (i.e. clustered together). Hence we use standard deviation σ to calculate the dispersion of the groups from the mean goal satisfaction. We consider the goal satisfaction quality to be

$$f_G(\alpha_k) = \frac{1}{1+\sigma_\alpha} \text{ where } \sigma_\alpha = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(f_g(g_i, \alpha_k) - \overline{f_g}\right)^2}$$

and the mean $\overline{f_g} = \frac{1}{N}\sum_{i=1}^{N} f_g(g_i, \alpha_k)$

4.4.2. Formation Quality
We refer by formation quality to how well were the groups formed in terms of satisfying *all the goals* for the collaboration task t.
• *Group Formation Quality:* This metric evaluates how well was a group formed in terms of all the goals. Similar to the previous calculation of group quality, for each group g_i

$$f_{fg}(g_i, t) = f_{fg}(g_i, \alpha_1, \alpha_2 \ldots \alpha_K) = \frac{1}{L}\sum_{j=1}^{L} f_g(g_i, \alpha_j)$$

• *Cohort Formation Quality:* This metric evaluates how well the cohort was formed in terms of all the goals and therefore the task. Similar to the previous calculations of cohort quality:

$$f_{fG}(t) = f_{fG}(\alpha_1, \alpha_2 \ldots \alpha_K) = \frac{1}{1+\sigma_f} \text{ where }$$

$$\sigma_f = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(f_{fg}(g_i, t) - \overline{f_{fg}}\right)^2} \text{ and } \overline{f_{fg}} = \frac{1}{N}\sum_{i=1}^{N} f_{fg}(g_i, t)$$

To analyze how useful (effective) are the constraints for a given goal, and the goals for a given task, we need to evaluate the formation quality of all possible formations over many runs using the same set of constraints for the goals. For each goal, if the resulted formation quality is constantly high, then if the goal satisfaction quality is high, and the constraint satisfaction quality for that goal is low, we consider that constraint to have a low significance in modeling that goal. Similarly, if the quality of the goal satisfaction is low, but the quality of constraint satisfaction for that goal is high, then the constraint has a low significance in modeling the goal. However, if the formation quality is low, then the constraint significance will be undefined despite the state of the goal satisfaction and the constraint satisfaction. For a large number of evaluated formations, we can evaluate the behaviour (*consistency*) and therefore the *reliability* of the constraints and goals, and consequently, the *effectiveness* of the formation using these constraints in the collaboration.

Optimal Formation: we define the optimal formation $form_{opt}(t)$ of the relation R as the optimal cohort that can result from the set of goals, such that the formation quality f_{fG} is maximized. We refer by $f_{optG}(t)$ to the quality of $form_{opt}(t)$.

So far, we assumed that to achieve the collaboration task, a CSGF system would apply the optimal formation to the given set of participants. However, unless the system is appointed to the optimal formation, it will select a formation at random. A possible way to know which formation is optimal is for the system to search for the optimal formation by calculating the quality of each possible formation generated by the set of given goals as shown in Procedure 1. These calculations however, mean that the system has to generate all the possible formations in order to return the

optimal one. Given the number of formed cohorts, the number of groups in each cohort, the number of goals, the number of constraints for each goal, the complexity of searching for the optimal solution is high.

Procedure 1 Calculating group formation quality
Given task t, $form_{opt}(t)$, $f_{optG}(t)$
  for each formation $form_n$
    for each group g_i in the cohort
      for each goal $\alpha_k \in t$
        for each constraint $c_i \in \alpha_i$
          calculate $f_{cj}(g_i, c_j)$
        calculate $f_g(g_i, \alpha_k)$
      calculate $f_{fg}(g_i, t) = f_{fg}(g_i, \alpha_1, \alpha_2 \ldots \alpha_K)$
    calculate $f_{fG}(t) = f_{fG}(\alpha_1, \alpha_2 \ldots \alpha_K)$
    if $f_{fG}(t) > f_{optG}(t)$
    then $form_{opt} \leftarrow form_n$
return $form_{opt}$

---

5. CONCLUSION AND FUTURE WORK
In this paper, we introduced a metrics framework for evaluating group formation based on constraint and goal satisfaction for CSGF systems within the leaning domain. As discussed before, the choice of constraints has a significant impact on the performance of the group. However, in this research, we consider the choice of constraints to be the responsibility of the formation initiator (in this case, the instructor) [4]. Therefore, a successful group is a group that has been well formed in relation to the constraints of the formation, and not one that performed well in the collaborative task. For future work, we intend to use this framework to evaluate the instructor based computer supported group formation system described in [4].

6. REFERENCES
[1] Bateman, B., Wilson, C. & Bingham, D. Team effectiveness development of an audit questionnaire. *Journal of Management Development,* Vol. 21 No 3**,** pp. 215-226, 2002.

[2] Higgs, M., Plewnia, U. & J. Ploch, J, Influence of Team Composition on Team Performance. Journal of *Team Performance Management,* Vol 11**,** pp 227-250, 2003

[3] McDonald, D. W. Evaluating Expertise Recommendations, In *Proceedings of ACM Group'01,* Boulder, Colorado, USA, pp 214-223, 2001

[4] Ounnas, A., H.C. Davis, and D.E. Millard, Towards Semantic Group formation. In *Proceedings of the* IEEE ICALT, Niigata, Japan, 2007.

[5] Redmond, M.A., A computer program to aid assignment of student project groups, Proc. of ACM SIGCSE, USA, 2001.

[6] Soh, L.-K., N. Khandaker, X. Liu, and H. Jiang. A Computer-Supported Cooperative Learning System with Multiagent Intelligence. In *Proceedings of* AAMAS'06, Hokkaido, Japan, 2006.

[7] Vivacqua, A. and H. Lieberman, Agents to assist in finding help. *In CHI*, Amsterdam, 2000, Vol 2, 1, pp. 65-72.

[8] Wessner, M. and Pfister, H. Group Formation in Computer-Supported Collaboration Learning. In *Proceedings of ACM Group'01.* Boulder, Colorado, USA, 2001, pp 24-31.