

Implementation of Group Formation Algorithms in the ELARS Recommender System

<https://doi.org/10.3991/ijet.v12.i11.7238>

Tina Knez^(✉), Martina Holenko Dlab, Natasa Hoic-Bozic
University of Rijeka, Rijeka, Croatia
t.inaknez92@gmail.com

Abstract—Collaborative learning is recognized as an effective way of gaining knowledge in an online environment. Therefore, e-courses frequently include collaborative e-learning activities (e-tivities) that are performed in pairs or small groups of students. One of the challenges for teachers who organize e-tivities is the effective group forming. This paper presents algorithms that can be used to divide a set of students participating in an e-tivity to homogeneous or heterogeneous groups. The criterion for automatic group formation includes the following characteristics: the program of study, gender, learning styles preferences, Web 2.0 tools preferences, knowledge level and activity level. Designed algorithms were implemented in the educational recommender system ELARS and tested in the context of e-tivities.

Key Words—Algorithms, collaborative learning, ELARS, group formation.

1 Introduction

According to the contemporary approaches to e-learning, students should not become passive recipients of information but active participants in the educational process [1], [2]. In order to place students in the center of the learning process, teachers design and conduct collaborative e-learning activities (called e-tivities [3]). E-tivities encourage interaction among students and empower them to take an active role in creation of their own knowledge. E-tivities can be performed in pairs, groups of three to six participants, and sometimes even in larger formations [2]. The formation of groups is an important prerequisite for carrying out collaborative e-tivities that will have positive effect on learning and intellectual growth. Inadequate group formation can demotivate students and cause problems such as disproportionate participation or resistance to collaboration in the future activities [4].

Formed groups can be homogeneous or heterogeneous. In a heterogeneous group members differ in their characteristics. Such groups are considered to be very efficient because they encourage interaction and students who achieve better results by helping their colleagues [5]. Homogeneous groups consist of the students with the same or similar characteristics. It can be useful to form homogeneous groups when the design of e tivity is tailored to certain knowledge level, learning style or similar [6].

Automatic group formation is often used in e-learning systems when a large number of students should be divided into groups based on a set of characteristics (attributes) [4], [5]. This paper aims to contribute to the field by presenting newly developed functionality of the proprietary e learning recommender system ELARS that tends to assist teacher in organizing e-tivities by providing automatic group formation support. The paper presents algorithms for dividing students to homogeneous or heterogeneous groups, designed in accordance with the data regarding students' characteristics available in the ELARS system. The paper also brings some implementation details and preliminary results of evaluation in the context of e-tivities.

2 Background

2.1 Grouping algorithms in e-learning systems

In the context of computer-supported collaborative learning, various approaches and algorithms are used to solve the problem of the automatic division of students into groups. In [4], authors gave an overview of proposed and implemented solutions showing that, in most cases, automatic grouping relies on probabilistic algorithms which employ a degree of randomness. In such approaches, the goal is not to find the best possible solution, but to find the approximation of the best solution [6], [7]. Examples are genetic algorithms [8] and those based on swarm intelligence [9]. Other approaches include semantic web, ontologies, Bayesian networks, and machine learning techniques [4].

In order to divide a set of students into groups using grouping algorithms, students' characteristics are usually represented as vectors [8]. There are examples where only one characteristic is used (e.g. learning styles [5]), but research in the field indicate that various attributes should be taken into account in order to maximize learning opportunities [8]. A set of characteristics can include intellectual abilities, gender, experiences, preferences, interests, personalities, etc. [7]. Once groups are automatically formed and the results are presented to the teacher, some systems allow the teacher to change the composition of the created groups or restart the grouping process in order to get different result [7]. Grouping algorithms are usually employed at the beginning of the e-tivity, but there are also approaches where they are used to rebuild groups during the e-tivity. For example, authors in [6] use the descent algorithm with the neighborhood of 2-opts to reconstruct the groups that are not homogeneous, according to the activity traces collected during the first part of the forum discussion.

2.2 Educational recommender system ELARS

The system ELARS is used to personalize e-learning experiences by determining recommendations to students regarding their participation in e-tivities. The system recommends optional e-tivities, possible collaborators, Web 2.0 tools, and provides advice regarding the level of participation [10], [11]. Using the ELARS web applica-

tion [12] teachers define course learning design with all planned activities and criteria for determining recommendations while students enter their preferences, read recommendations and choose among the recommended items. The main components of the system are *activity model* (describes learning design of an e-course), *student and group models* (represent student's and group's characteristics), and *recommender module* (contains rules and algorithms for generating recommendations).

The most important component for group formation is the student model as it contains students' characteristics used in grouping criterion. The following characteristics are represented as continuous variables with values ranging from -1 to 1 [10]: *preferences of Web 2.0 tools* and *VARK learning styles* (determined using questionnaires), *knowledge level* (calculated based on test results), and *activity level* (calculated based on activity data retrieved from Web 2.0 tools). The student model also contains two characteristics with nominal values: *gender* and *program of study*.

3 Automatic group formation

Up to now, the ELARS system encouraged grouping for an e-tivity in line with the criterion set by a teacher but students were not obliged to choose recommended colleagues. The aim of the implementation of the automatic group formation is to support teacher in dividing the students enrolled in the e-tivity into groups according to the specific criterion. This was achieved by implementing group formation algorithms and recommending grouping of the students to the teacher.

During the development of the automatic group support, the goal was to provide the teacher with flexibility in defining the grouping criterion. Therefore, the teacher is empowered to define the grouping criterion by combining numerical and nominal characteristics from the student model. Additionally, depending on the design of the e-tivity, he/she can decide whether to form homogenous or heterogeneous groups. The algorithms for forming homogeneous and heterogeneous groups were designed according to the approach which does not tend to determine the best possible solution but uses heuristics to approximate a good solution, as proposed in [7].

3.1 Grouping criterion

Let S be the set of students who participate in an e-tivity, $|S| = n$. To divide students from S in groups that will have about the same number of members, teacher adjusts the grouping criterion in line with the designed e-tivity. The criterion includes a set of variables: maximum number of students in the group M , grouping method gm (homogeneous or heterogeneous groups), and weighting factors w_i for characteristics from the student model. The sum of weights w_i assigned to the characteristics represented as continuous variables should be equal to 1, $w_i \in [0,1]$. In case of nominal characteristics it should be $w_i \in \{0,1\}$. All characteristics c_i , $i=1, \dots, t$, for which $w_i > 0$ will be taken into account during group formation. Students can be grouped based on a single characteristic (e.g. preferences of Web 2.0 tools offered for the realization of an e-tivity) or based on a subset of characteristics (e.g. combination of gender, learning style and

knowledge level). By setting the weighting factors in the grouping criterion, teacher specifies to what extent certain characteristic should affect the grouping.

A set of the following variables is included in the grouping criterion as well: distance between a pair of students (d), group average distance (gad), internal distance (igd), pair threshold (pt), and group threshold (gt). The *distance between a pair of students*, s_j and s_k , is calculated using Euclidean distance between the vectors representing their characteristics (the similarity between students is higher if d is closer to zero). *Internal group distance* is calculated as a sum of distances between pairs of group members using expression (1) where t is the number of characteristics. *Group average distance* is calculated as arithmetical mean of distances between pairs of group members [7].

$$igd = \sum_{j=1}^{p-1} \sum_{k=j+1}^p d(s_j, s_k) = \sum_{j=1}^{p-1} \sum_{k=j+1}^p \sqrt{\sum_{i=1}^t w_i (s_{ji} - s_{ki})^2} \quad (1)$$

To ensure homogeneity or heterogeneity of the groups, variables pt i gt are used in grouping algorithms as *allocation criterion* – to decide whether a student should be allocated to a certain group [7][1]. The variable pt represents the upper limit for the distance between students in the case of homogeneous grouping (the initial value of pt is set to 0.1) and lower limit in the case of heterogeneous grouping (the initial value of pt is set to 0.3). The value of gt is not set to the initial value but calculated depending on the number of group members, m , using (2). It represents the upper limit for internal group distance in the case of homogeneous grouping and lower limit in the case of heterogeneous grouping. To allow the algorithm to randomly allocate certain number of students to groups, *random grouping criterion* is set to 20%, meaning that if there is less than 20% percent of ungrouped students, they will be randomly allocated to the unfilled groups. In case students' characteristics are such that the allocation criterion cannot be satisfied, allocation threshold pt is automatically increased/decreased by 0.025. The initial values of pt (0.1, 0.3), *increase/decrease step* (0.025) and *random grouping criterion* (20%) were determined based on the experimental results.

$$gt = pt \cdot (m - 1) \quad (2)$$

3.2 Algorithm 1 – Calculating group sizes

The algorithm determines a list G with group sizes g depending on the number of students, n , and the maximum number of students per group, M . Group sizes are calculated in a way that the total number of groups is minimal. The following steps are used:

STEP 1: Calculate the required number of groups x , $x = \text{ceil}(n/M)$.

STEP 2: Calculate the group size g , $g = \text{ceil}(n/x)$; add g to the list G ; reduce n by g , and reduce x by 1. This step is repeated until $n > 0$.

3.3 Algorithm 2 – Forming homogeneous groups

The main steps of the algorithm for forming a list of homogeneous groups are:

STEP 1: For each group, select a random student-pivot and add it to the group.

STEP 2: For each unfilled group, try to fill the group with ungrouped students who meet the allocation criterion. To check whether a student is a candidate for entering an unfilled group, a comparison of student's and pivot's nominal characteristics for which $w_i=1$ is performed. If those characteristics match, and the allocation thresholds pt and gt are satisfied, the student is allocated to the group.

STEP 3: If there is a group with less than three members, cancel the current grouping, and try again from Step 1. If retrieval fails, increase pt for 0.025 and start all over again. The impossibility to form groups indicates that pt and gt values are set too strictly so they are automatically increased before repeating the procedure.

STEP 4: Try to fill unfilled groups by replacing the pivots. Some of the ungrouped students might enter the unfilled groups based on the higher similarity with the new pivots.

STEP 5: If random grouping criterion is satisfied, go to Step 6. Otherwise, cancel the current grouping, increase pt for 0.025 and start all over again from Step 1.

STEP 6: For each unfilled group, try to add ungrouped students who meet the allocation criterion to the group. If such do not exist, fill groups with random students.

3.4 Algorithm 3 – Forming heterogeneous groups

This algorithm determines a list of heterogeneous groups in the following way:

STEP 1: For each group, select a student-pivot and add it to the group: in case there are no nominal characteristics (gender, study program) with $w_i=1$, select a random student-pivot. Otherwise, select a random student-pivot whose gender and/or study program are in the minority.

STEP 2: For each unfilled group, try to fill the group with ungrouped students who meet the allocation criterion. The candidates for entering an unfilled group are students of opposite gender and/or different study program compared to the gender/study program of the student who had previously entered the group.

STEP 3: If there is a group with less than three members, cancel the current grouping, try again from Step 1. If retrieval fails, decrease pt for 0.025 and start all over again from Step 1.

STEP 4: In case there are nominal characteristics with $w_i=1$, try to fill each unfilled group with ungrouped students whose characteristics satisfy the thresholds pt and gt (regardless the gender and/or study program). Otherwise, skip this step.

STEP 5: Try to fill unfilled groups by replacing the pivots. It is possible that some of the ungrouped students will be less similar to the new pivot than to the old pivot.

STEP 6: If random grouping criterion is satisfied, go to Step 7. Otherwise, cancel the current grouping, decrease pt for 0.025 and start all over again from Step 1.

STEP 7: For each unfilled group, try to add ungrouped students of opposite gender and/or different study program compared to gender/study program of the student who had previously entered the group. If such do not exist, fill groups with random students.

4 Implementation and preliminary results

The process of automatic group formation, shown on Fig. 1, is implemented within the recommender module of the ELARS system. Using the ELARS web application, teacher should define grouping criterion and start the process. After retrieving grouping criterion and students' data, the system will determine group sizes, execute appropriate grouping algorithm, and recommend grouping to the teacher. In the list of the proposed groups (Fig. 2), students' characteristics are shown and teacher can examine the composition of a particular group. Based on the displayed *gad* value (group average distance), he/she can compare to what extent are the groups homogeneous or heterogeneous (a value closer to zero indicates more homogeneous group). Upon receiving the recommendation, teacher has the opportunity to save the formed groups or to restart the process with the same or revised grouping criterion. Algorithms include a certain level of randomness so different runs will produce different results. This enables the teacher to select grouping that suits the best to the designed e-tivity.

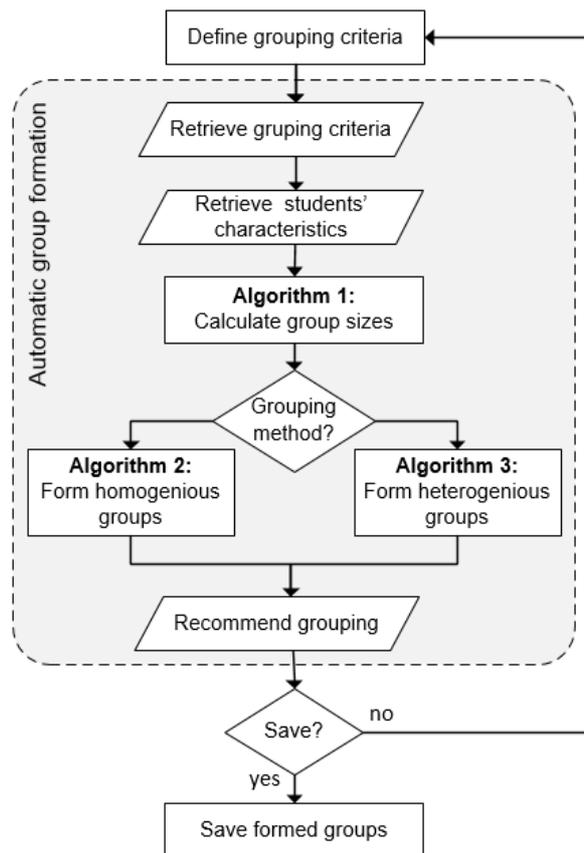


Fig. 1. Group formation process in the ELARS system

In order to test the implemented functionality, an e-tivity was designed and performed during two computer science courses with 26 and 40 enrolled students. During the period of three weeks, groups of 3-5 students were supposed to write an analysis of examples related to the given topic using Wikispaces [2], a collaborative writing platform. Students could earn 20 out of 100 course points for this e-tivity, depending on the quality/quantity of contribution. Grouping criterion was defined to form heterogeneous groups in respect to the students' knowledge level (calculated based on the online test results for the corresponding lessons), and the preferences of the Wikispaces and VARK learning styles. The objective was to group together the students with different preferences as well as different knowledge of the subject they are writing about. In that way, students with higher knowledge level were given the opportunity to correct group members' mistakes or omissions. In both courses, groups were formed using proposed algorithms and all groups successfully finished the e-tivity (the average result was above 80%).

Analysis of contributions showed that students, besides writing their own contributions, tried to improve the text written by their group members in order to achieve better final result. This indicates that groups formed by ELARS really collaborated in accordance with the aim of the designed e-tivity. The results of the survey performed with students (N=64) show that 80% of the students think it is effective that e-tivity was performed in groups. Most students consider that groups should be created based on the students' characteristics (47% of students) and do not support entirely random grouping (53% of students). In their comments, many students pointed out that they would prefer the ability to create groups by themselves and stated that they would always choose the same collaborators - colleagues with whom they previously had a successful collaboration. On the other hand, 80% of the students are well aware that through collaboration with different colleagues they have chance to acquire teamwork skills needed for their future professional work. These finding indicates the need for combining both teacher's criterion and students' preferences in the group definition process.

According to teachers' comments, the implemented group formation support is very useful. However, they suggested improvements of the graphical representation of the recommended grouping to make it even clearer whether the created groups meet the defined criteria.

5 Conclusions

The presented research aims to support teachers in the process of organizing e-tivities by providing automatic group formation support. To achieve that, the recommender module of the ELARS system was supplemented with the algorithms for dividing students into homogeneous or heterogeneous groups. Teacher is empowered to choose grouping method adequate to the planned e-tivity and define to what extent particular characteristic should affect the grouping. Preliminary results of using proposed algorithms were obtained in the context of several short-term e-tivities.

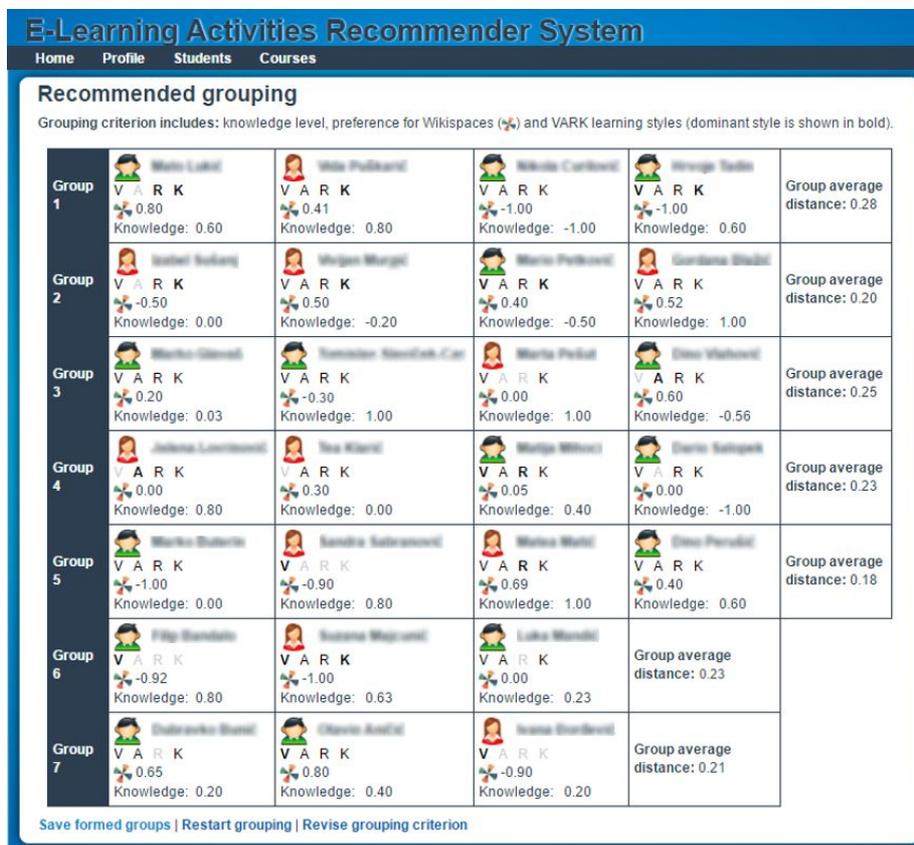


Fig. 2. Grouping recommendation (names of the students are blurred)

Insight into students' activity in Wikispaces and results of surveys conducted among teachers and students confirmed that students achieved a successful collaboration. The teachers consider implemented automatic group formation to be useful but suggest improvements of the graphical representation of the students' characteristics - members of the formed groups. Students are aware of the benefits of group forming based on the certain characteristics but would also like to collaborate with colleagues with whom they have previously had positive experiences in collaborative e-tivities. These issues will be addressed during further development of the automatic group support in the ELARS system. Besides using visualization techniques for presenting recommended grouping, the student model will be extended with data regarding students' preferences towards other students. By providing this information, the student will be able to express to what extent is a particular colleague his/her desirable collaborator for realization of collaborative e-tivities. The inclusion of these data in the grouping algorithms will enable the automatic creation of groups that meet the pedagogic criterion set by the teacher as well as the students' preferences.

Further work will also enable the automatic creation of mixed groups that satisfy heterogeneity for one subset of characteristics and homogeneity for another subset of characteristics. Finally, further research will include a more extensive evaluation in order to determine the effect of the implemented algorithms on students' level of engagement as well as their academic results during various types of e-tivities (e.g. long-term e-tivities designed according to the problem-based learning strategy).

6 Acknowledgment

The research has been conducted under the project “E-learning Recommender System” (reference number 13.13.1.3.05) supported by University of Rijeka (Croatia).

7 References

- [1] B. Aguti, R. Walters, and G. Wills, “Effective use of e-learning technologies to promote student-centered learning paradigms within higher education institutions,” *Int. J. e-Learning*, pp. 1–8, 2014.
- [2] H. Beetham, “An approach to learning activity design,” in *Rethinking Pedagogy for a Digital Age: Designing for 21st century learning*, 2nd ed., H. Beetham and R. Sharpe, Eds. Routledge, 2013, pp. 31–48.
- [3] G. Salmon, *E-tivities: the key to active online learning*. Psychology Press, 2002.
- [4] W. M. Cruz and S. Isotani, “Group Formation Algorithms in Collaborative Learning Contexts: A Systematic Mapping of the Literature,” in *Proceedings of Workshop on CYTED-RITOS International*, 2014, pp. 199–214.
- [5] A. Gogoulou, E. Gouli, G. Boas, E. Liakou, and M. Grigoriadou, “Forming homogeneous, heterogeneous and mixed groups of learners,” *Proceedings of Workshop on Personalisation in E-Learning Environments at Individual and Group Level*, 2007, pp. 33–40.
- [6] M. Salihoun, F. Guerouate, N. Berbiche, and M. Sbihi, “How to Assist Tutors to Rebuild Groups Within an ITS by Exploiting Traces. Case of a Closed Forum,” *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 3, pp. 169–181, 2017. <https://doi.org/10.3991/ijet.v12i03.6506>
- [7] P. Paredes and A. Ortigosa, “TOGETHER: An authoring tool for group formation based on learning styles,” in *Proc. A3H: 7th international workshop on authoring of adaptive and adaptable hypermedia at EC-TEL 2009*, 2009.
- [8] A. Sukstrienwong, “A Genetic-algorithm Approach for Balancing Learning Styles and Academic Attributes in Heterogeneous Grouping of Students,” *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 3, pp. 4–25, 2017. <https://doi.org/10.3991/ijet.v12i03.5803>
- [9] H. Alhunitah and M. E. B. Menai, “Solving the student grouping problem in e-learning systems using swarm intelligence metaheuristics,” *Comput. Appl. Eng. Educ.*, vol. 24, no. 6, pp. 831–842, Nov. 2016. <https://doi.org/10.1002/cae.21752>
- [10] N. Hoic-Bozic, M. Holenko Dlab, and V. Mornar, “Recommender System and Web 2.0 Tools to Enhance a Blended Learning Model,” *IEEE Trans. Educ.*, vol. 59, no. 1, pp. 39–44, 2016. <https://doi.org/10.1109/TE.2015.2427116>
- [11] M. Holenko Dlab and N. Hoic-Bozic, “Increasing students' academic results in e-course using educational recommendation strategy,” in *Proceedings of the 17th International Con-*

ference on Computer Systems and Technologies 2016 - CompSysTech '16, 2016, pp. 391–398.

[12] “ELARS,” (in Croatian), 2015. [Online]. Available: <http://elars.uniri.hr/elars>. [01-Feb-2017].

[13] “Wikispaces,” 2015. [Online]. Available: <https://www.wikispaces.com/>. [01-May-2017].

8 Authors

Tina Knez finished her master study at Department of Informatics, University of Rijeka, HR-51300 Rijeka, Croatia (tinaknez92@gmail.com).

Martina Holenko Dlab is a postdoctoral researcher at Department of Informatics, University of Rijeka, HR-51300 Rijeka, Croatia (mholenko@inf.uniri.hr).

Natasa Hoic-Bozic is a professor at Department of Informatics, University of Rijeka, HR-51300 Rijeka, Croatia (natasah@inf.uniri.hr).

Article submitted 1 June 2017. Published as resubmitted by the authors 11 September 2017.