A Group Formation Tool in a E-Learning Context

Christos E. Christodoulopoulos
College of Science and Engineering
School of Informatics
University of Edinburgh, UK
christos.c@ieee.org

Kyparisia A. Papanikolaou
General Department of Education
School of Pedagogical and
Technological Education, Greece
spap@di.uoa.gr

Abstract

In this paper we present a web-based group formation tool that supports the instructor to automatically create both homogeneous and heterogeneous groups based on up to three criteria and the learner to negotiate the grouping. Moreover, the instructor is allowed to manually group learners based on specific criteria. A discriminative feature of this tool is the use of the Fuzzy C-Means algorithm for homogeneous grouping, which provides for each learner the probability of belonging to different groups. This information is also provided to the instructor to support him/her in manually exchanging learners or intervening in the initial grouping. Moreover, the learners are informed for the groups formed and they are allowed to negotiate their group assignment. Preliminary evaluation results provide indications for the efficiency of the proposed approach in forming homogeneous and heterogeneous groups in a real context.

1. Introduction

Group work, under proper conditions, encourages peer learning and support providing an opportunity for students to clarify and refine their understanding of concepts through discussion and rehearsal with peers. However, several factors have been investigated for their influence on the group dynamics and performance, such as group synthesis. It is proposed that variation in group inputs such as member abilities/skills/relations and group structure, fosters different types of interaction and outcomes [7].

We regard grouping as important for student learning mainly because of its shaping force on instruction and students' social participation. Moreover, as critical factors in forming groups we consider learners' individual characteristics/attributes relative to learning such as knowledge and learning style.

A variety of group formation techniques have been used in forming learning groups such as random assignment, learner-formed groups or grouping according to academic (e.g. knowledge of a subject), social (e.g. gender), traits (e.g. learning style) or learner context. However, the relative merits of homogenous and heterogeneous groups seem to depend on several factors such as students' abilities, traits, curriculum area, and task [14, 9, 5].

Research in the field of e-learning has lead to the development of computer-based tools that support automatic group formation. Learners are dynamically assigned to groups according to their learning needs or individual characteristics or they are matched to peer learners for a specific task upon their request [9, 6, 8]. For instance in the DIANA system [12], the instructor can perform homogeneous groupings by choosing multiple criteria (up to 7) but the system does not support manual group formation, or the editing of the created groups. In the system proposed by Cavanaugh et. al [2], the instructor may select multiple criteria assigning weights to each one, whilst surveys are used to automatically identify learners' personality traits. This system uses the Hill Climbing algorithm for both homogeneous and heterogeneous groupings. In the OmadoGenesis tool [4], the instructor is allowed to (a) select the set of learners to apply the grouping, (b) select the grouping type (homogeneous or heterogeneous) for each of the criteria used (up to 3), and (c) edit/reform the created groups. However the instructor cannot manually group students but s/he is allowed to intervene in the proposed grouping exchanging learners among groups. This tool forms homogeneous groups using the k-means clustering algorithm, heterogeneous using a heterogeneity matrix and mixed type groups using Genetic Algorithms. The use of Genetic Algorithms is also proposed by Wang et. al [12], but only for creating heterogeneous groups. Useful options provided by the tool of Cavanaugh et. al [2] and OmadoGenesis [4] to the instructor are to manually define or reform the groups respectively. However none of the above systems support the instructor in this process by providing advice about how to group students or interchange them among the groups. Moreover, all the above systems seem to focus on the instructor as none

of them provides the learner with access to the grouping process.

In this paper we present a web-based group formation tool that can be used as a stand-alone web-based application, or as a module of an e-learning environment for matching peers based on specific criteria. The tool supports the instructor to create both homogeneous and heterogeneous groups based on up to three criteria and the learner to negotiate the grouping. Moreover, the instructor is allowed to manually group students based on specific criteria. A discriminative feature of this tool is the use of the Fuzzy C-Means (FCM) algorithm for creating homogeneous groupings. The main advantages of FCM are (i) its ability to work in spaces that contain a limited amount of data (i.e. students in class 20-100) and with small groups, (ii) the membership function, i.e. in FCM a learner may belong to more than one groups with a different probability. This information is particular useful for the instructor when he needs to manually exchange learners. Using this information, the instructor can easily identify the groups that a learner may belong to based on his/her individual characteristics without having to cross-check all the different groups and compare their characteristics to the learner profile. Moreover, the learners are informed for the grouping and they are allowed to negotiate their group assignment.

The paper is structured as follows. In section 2 the algorithms used for homogenous and heterogenous groupings are described. Following, in section 3 we focus on the design and the implementation of the tool. Then in section 4 preliminary evaluation results are presented. Finally in section 5, conclusions and further research plans are presented.

2. Algorithms for grouping learners

Research on algorithms adequate for learning group formation has resulted to a limited number of proposals. Specifically the majority of the related studies focus on optimization algorithms such as Hill Climbing, Ant Colony Optimization or Genetic Algorithms. While such algorithms tend to perform adequately for either homogeneous, heterogeneous or mixed-type (i.e. homogeneous for one attribute and heterogeneous for another) groupings, their computational complexity (polynomial) is high and implementation is quite hard for purely web-based environments that use "simple" scripting languages such as JavaScript and PHP.

With the low complexity parameter as a guideline we turned to clustering algorithms and more particularly to the c-means algorithms family. In a previous study described in [3] we compared the two well-known c-means clustering algorithms; k-means and Fuzzy C-Means (FCM). They both present linear complexity $O(n)$ [11] and various other features that make them promising candidates for supporting homogeneous groupings. In particular, both algorithms

in their standard forms were evaluated in a simulated environment proving that the "fuzzy" nature of the FCM is a promising alternative for creating groups.

Regarding heterogeneous grouping, research results have been very limited until now. Alternative definitions have been proposed for heterogeneity. According to Graf and Bekele [5]:

> A reasonably heterogeneous group refers to a group where student-scores reveal a combination of low, average and high student-scores.

However this approach is limited to 3 discrete classes of attribute values (student-scores). Heterogeneity may also refer to specific group attributes or to the group as a whole. In our case, extending the previous definition, we define a heterogenous group as a group where all the possible values of the learner space (involving all the different group attributes) are present. As discussed in the following section, the use of a random algorithm seems adequate for creating heterogeneous groups with satisfying results.

The learner space model is multi-dimensional; each dimension represents a learner attribute (i.e. learner characteristic) used for group formation purposes. Therefore every learner j is represented by a vector $\vec{x}_j = [d_1, d_2, \ldots, d_n]$ where d_z is the value of the z_{th} attribute. The values of each attribute have to be integers with no restriction about their range; it is acceptable for the space model to have uneven dimensions.

For instance, $x_3 = [9, -11, 2]$ represents the vector for the learner with id number 3, that has values $d_1 = 9$, $d_2 = -11$ and $d_3 = 2$ for attributes 1,2 and 3 respectively.

FCM and random sorting algorithms are able to work on n-Dimensional spaces. In the current implementation of the tool we allow the use of maximum 3 attributes as grouping criteria in order to simplify the grouping process and facilitate instructors in intervening to this process when reviewing the group formation results.

2.1. Homogeneity based on FCM

Fuzzy C-Means (FCM) or Fuzzy ISODATA algorithm was proposed by James Bezdek in 1981 [1] and since then it evolved into one of the most widely used fuzzy clustering algorithms in data clustering, pattern recognition and image processing with very satisfying results [10, 13].

FCM is essentially the fuzzy version of the well-known k-means algorithm and thus their structure is very similar. However in FCM a data point may belong to more than one cluster with a specific membership probability for each cluster. So for every cluster a membership matrix (U) is created to represent the membership probabilities for every data point.

Clustering algorithms are special forms of optimization algorithms as they can only minimize an objective function. In the case of FCM that function is described in Eq. 1.

$$J = \sum_{i=1}^{c} \sum_{j=1}^{n} \left(u_{ij}^m d(\vec{c_i}, \vec{x_j}) \right) \quad (1)$$

where i is the current cluster,
$\vec{x_j}$ is the current data point,
$\vec{c_i}$ is the center of cluster i,
u_{ij} is the membership probability of $\vec{x_j}$ for cluster i,
$d(\vec{c_i}, \vec{x_j})$ is the Euclidian distance between cluster center $\vec{c_i}$ and data point $\vec{x_j}$ and
m is a weighting exponent with $m > 1$
note that both cluster centers and data points are represented by vectors as the learner space is multi-dimensional; every dimension represents one different learner attribute.

The current implementation of the FCM algorithm in the Group Formation Tool, is described below:

1. Initialize Probability Matrix U:

    (a) Assign random values to U

    (b) For every member j, make u_j sums to 1
    $u_{ij} = u_{ij} / \sum_{j=1}^{n}(u_j)$

2. For every step κ:

    (a) Calculate new cluster centers for each cluster i:

    $$c_i = \frac{\sum_{j=1}^{n}(u_{ij}^m \cdot \vec{x_j})}{\sum_{j=1}^{n}(u_{ij}^m)}$$

    (b) Calculate the distance matrix for every dimension z:

    $$d_{ij} = \sqrt{\sum_{z=1}^{d}\left((x_{jz} - c_{iz})^2\right)}$$

    (c) Calculate the new probability matrix:
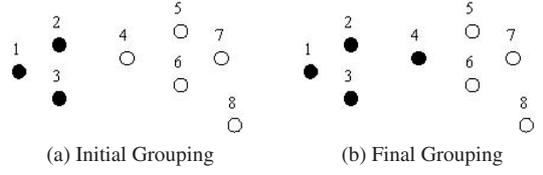
    $$u_{ij} = \frac{d_{ij}^{-\frac{2}{m-1}}}{\sum_{i=1}^{c}(d_{ij})}$$

    (d) Calculate the value of the objective function:

    $$J^{\kappa} = \sum_{i=1}^{c} \sum_{j=1}^{n} \left(u_{ij}^m d(\vec{c_i}, \vec{x_j}) \right)$$

    (e) If $\|J^{\kappa} - J^{\kappa-1}\| < min\_impro$ then end

    where $0 < min\_impro < 1$ is the termination criterion

(a) Initial Grouping   (b) Final Grouping

$$U_1 = \begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0.9871 & 0.9745 & 0.9832 & 0.4364 \\ u_{15} & u_{16} & u_{17} & u_{18} \\ 0.2634 & 0.1485 & 0.0847 & 0.0325 \end{bmatrix}$$

$$U_2 = \begin{bmatrix} u_{21} & u_{22} & u_{23} & u_{24} \\ 0.0129 & 0.0255 & 0.0168 & 0.5636 \\ u_{25} & u_{26} & u_{27} & u_{28} \\ 0.7366 & 0.8515 & 0.9153 & 0.9675 \end{bmatrix}$$

(c) Membership Matrix

Figure 1: Exchanging data points using the FCM Membership Probability Matrix

The main problem that we encountered with FCM in grouping learners (as also with the k-means algorithm) was the inequality of the clusters created. The source of this problem is that both algorithms take as input the desired number of clusters and not the number of data points per cluster (i.e. learners per group). During the clustering process, data points with extreme values tend to isolate; in such cases some of the created clusters might have significantly less members than others. Thus, the challenge here is to amend this situation, i.e. automatically reform the clusters in order to become of equal size, preserving the nature of the groups (homogenous). To this end, we developed an algorithm utilizing the FCM's probability matrix which proved extremely useful for interchanging data points among clusters of different size. In particular, aiming to form groups of equal size with the best possible clustering quality (although lower than the initial), we exchange members among the groups based on their highest probability of belonging to particular clusters as this is proposed by the FCM. Let's illustrate the proposed approach through an example. Suppose that we have 8 data points that FCM grouped in two clusters as illustrated in Fig. 1a: the three black bullets belong in cluster 1, whilst the five white ones in cluster 2. Outputs of FCM are the two matrices U_1 and U_2 (Fig. 1c) that contain the membership probabilities of the data points with respect to cluster 1 and 2. At the end of the grouping process, the data points represented in U_1 with higher probabilities are assigned to cluster 1, whilst those of U_2 with higher probabilities are assigned to cluster 2. The problem is that cluster 1 (3 data points) has one member less than cluster 2 (5 data points). However, the data point that corresponds

to the [1,4] position of U_1 and U_2 (see Fig. 1c), is the best possible exchange that can be made between the clusters. That is because the data point U_1 [1,4] has the highest probability among the data points of U_1 that weren't included in cluster 1 (i.e. 0.4364, 0.2634, 0.1485, 0.0847, 0.0325), and at the same time the data point U_2 [1,4] has the lowest of those included in cluster 2 (i.e. 0.5636, 0.7366, 0.8515, 0.9153, 0.9675). Therefore, in order to evenly distribute the 8 data points into two clusters, we decided to transfer data point [1,4] to cluster 1 as illustrated in Fig. 1b.

The above approach enabled us to reduce the required amount of procedures to just basic conditional search implemented through the following "equalization" function:

1. Calculate the number of members per group:

$$n = (round(total\_students/groups))$$

where n, is the number of members per group $total\_students$, is the total number of students and $groups$, is the number of groups to be created

2. For every cluster i

   (a) Find if there are any extra members:

   $$\text{if } count(c_i) > n \text{ then}$$

   (b) For every extra member j that has the lowest probability for this cluster:

   $$\text{for } j \text{ where } u_{ij} = \min\{u_i\}$$

   i. Search for the cluster with the best probability (besides the current one):

   $$\text{find } x \text{ where } u_{xj} = \max\{ui\} \text{ and } x \neq i$$

   ii. If the new cluster x has the maximum allowed number of members repeat the previous step with $x' \neq x$ and $x' \neq i$

In the current implementation of the Group Formation Tool, the instructor is offered the option to activate the equalization function and automatically redistribute learners to groups, or do it manually taking into account the information provided about their group membership probability.

2.2. Heterogeneity based on Random Algorithm

To address the issue of heterogeneity we propose the use of a standard random algorithm; that is to apply a uniform distribution on the learner space. The core idea of this approach is based on the nature of heterogeneity. In our case, a heterogeneous group is considered as a group that contains all the possible values of the learner space. In a real

Group	Name	FCM Membership Probabilities	Knowledge Level	Felder-Silverman LSI	LSQ Reflective - Active	LSQ Pragmatist - Theorist
title: Intelligent and Adaptive Learning Environments						
Group 1						
1	?????? ????????	Group 1: 99.86% Group 2: 0.03% Group 3: 0.05% Group 4: 0.04% Group 5: 0.02%	2	Strong Global Very Strong Sensing Moderate Visual Moderate Active	Low Active	Moderate Theorist
4	???????????? ?????????	Group 1: 20.37% Group 2: 7.34% Group 3: 10.19% Group 4: 54.15% Group 5: 7.95%	1	Very Strong Global Moderate Sensing Moderate Verbal Very Low Reflective	Moderate Reflective	Low Pragmatist
1	??????? ???????????	Group 1: 99.86% Group 2: 0.03% Group 3: 0.05% Group 4: 0.04% Group 5: 0.02%	3	Moderate Global Very Strong Sensing Moderate Visual Moderate Reflective	Low Active	Moderate Theorist

Figure 2: A screen shot of the Group Formation Tool (Instructor Environment: Edit mode). The learners' names have been deliberately replaced by question marks "?"

context where it is not possible to cover the whole spectrum of the space, heterogeneous groups would contain all possible values based on the available data. Assuming that all the values should have the same possibility to belong to a certain group, we decided to use the uniform distribution for forming heterogeneous groups. In particular, a random selection algorithm is used for forming groups that their members' attributes span across the learner space. This algorithm provides an adequate level of heterogeneity while being extremely fast and easy to implement [3].

Especially, we do not use random assignment rather we propose a selection (without replacement) that follows a uniform distribution over specific criteria. However the efficiency of this algorithm needs to be evaluated and compared to other algorithms used for heterogeneous group formation in real-life environments.

3. The Group Formation Tool

The Group Formation Tool was developed in PHP. It uses a MySQL database built to support multiple lessons/instructors. The algorithms used for homogeneous/heterogeneous group formation were implemented in a modular way, allowing ease integration of the tool in any e-learning environment.

The tool allows the instructor to form groups either manually or automatically. The instructor is able to select the grouping method (manual or automatic) and then define the number of groups to be created. In case of the automatic group formation, the instructor selects the type of grouping (homogeneous or heterogeneous) and up to three differ-

Figure 3: A screen shot of Group Formation Tool (Learner Environment: Group Negotiation). The learner is allowed to negotiate his re-assignment to a different group through the 'Group Negotiation' form. The name of the learner has been deliberately replaced by question marks "?"

ent grouping criteria (e.g. knowledge level, learning style). Finally the instructor is allowed to activate the equalization function ensuring the equal distribution of learners to groups.

After the automatic group formation process, the resulted groups are presented to the instructor in the form of separate tables. For each learner, the system presents his/her name and particular characteristics. The Group Formation Tool is able to personalize all the group listings allowing instructors to select the set of individual learner characteristics to appear.

The instructor has always the option to edit the output of the group formation process (i.e. re-assign the learners to different groups) by using the Edit mode. In this mode, a more detailed profile for each learner appears. In case of homogeneous grouping, the learner profile appears along with his/her Membership Probabilities for each group. This feature is very useful in case the instructor decides to check the best possible grouping using the FCM and then to manually perform the equalization of the number of members per group. Another scenario is that the instructor needs to intervene in the grouping process as s/he is aware of a potential rivalry between two group members, or a specific request of a learner for re-assignment.

A screenshot of the Edit mode is presented in Fig. 2. In this figure the profiles of three members of Group 1 (data represented in Table 1) are shown including the FCM generated membership probabilities, knowledge level for the current lesson, style according to the Felder-Silverman model, the Reflective/Active and the Pragmatist/ Theorist axis of the Honey & Mumford model. Let's assume that the instructor wants to re-assign one learner from Group 1 to another group. To achieve that, s/he must first pick the learner that has the least membership probability of belonging to Group 1. In this case the second learner (20.37% membership probability) is the most appropriate candidate. Then the instructor should find the best possible group to assign the particular learner (in a similar way with the equalization function described above): the group that the particular learner has the highest membership probability to belong to, i.e. Group 4 (54.15%). The instructor then simply replaces the group number in front of the learner (see Fig. 2) and saves the new group. This method seems more accurate and time-saving compared to the cross-checking of every attribute value of the learner with those of learners of other groups.

The Group Formation Tool provides also a number of useful functions to learners. Learners can view their profile and alter it. In this case the instructor will be informed for any changes when s/he visits the particular learner profile. A learner is also able to view his/her group assignment per course along with all the group members. Finally the learner is facilitated to negotiate his/her assignment with the instructor in case s/he wants to. Fig. 3 illustrates this option. The learner is allowed to submit a form to the instructor stating the reasons for his/her request of reassignment.

4. Evaluation

In order to evaluate the efficiency of the proposed approach for homogenous and heterogeneous groupings (the FCM algorithm combined with the equalization function and the Random Selection Algorithm), we performed a series of tests using real learner data. In particular, we used data from 18 undergraduate students (4th year students) of the Department Technology Education and Digital Systems of the University of Piraeus who had enrolled in the module "Intelligent and Adaptive Learning Environments" (academic year 2005-2006).

Below we provide two different groupings, homogeneous and heterogeneous produced by the Group Formation Tool. The results collected from the instructor environment of the tool are presented in Tables 1 and 2 for illustration purposes. Table 1 represents the homogeneous grouping whilst Table 2 the heterogeneous grouping.

Homogeneous grouping. In the first test we aim to assign the 18 learners to 5 homogeneous groups using two criteria. The first criterion was the Sensing/Intuiting axis of the Felder & Silverman learning styles model and the second was the Reflective/Active axis of the Honey & Mumford model. Both criteria are represented by 6 discrete values for each side of the axis, negative ones correspond to the left side of the axis, e.g. Sensing for the Sensing/Intuiting axis, whilst positive values to the right side, e.g. Intuiting for the Sensing/Intuiting axis.

Table 1: Test group formation using FCM and equalization function. All the membership probabilities for Learner no.2 appear as s/he was automatically re-assigned by the tool

Learner	Group	Membership Probability	Sensing/ Intuiting	Reflective/ Active
1	1	99.86%	-11	3
2	1	20.37%	-5	-5
	2	7.34%		
	3	10.19%		
	4	54.15%		
	5	7.95%		
3	1	99.86%	-11	3
4	2	91.77%	11	3
5	2	93.84%	9	1
6	2	71.57%	7	1
7	2	91.77%	11	3
8	3	62.06%	7	9
9	3	62.02%	-1	7
10	3	53.55%	3	1
11	3	94.05%	5	5
12	4	63.04%	3	11
13	4	80.90%	-5	11
14	4	63.04%	3	11
15	4	99.23%	-1	-9
16	5	98.02%	11	-9
17	5	98.02%	11	-9
18	5	97.69%	9	-9

1(-1)	Very Low
3(-3)	Low
5(-5)	Mediocre
7(-7)	Moderate
9(-9)	Strong
11(-11)	Very Strong

For instance learner no. 1 is characterized as a 'Very Strong Sensing' and 'Low Active' (values -11 and 3 respectively).

Preliminary results indicate the potential of the proposed approach for homogenous groups. In particular, the 18 learners have been evenly distributed in five groups; 3 groups consisting of 4 learners and 2 group of 3 learners. Especially, as illustrated in Table 1 groups 2, 3 and 5 comprise of learners with very close values based on both criteria. Groups 1 and 4 present an adequate level of homogeneity although learners 2 and 15 seem not to fit very well within these groups; their Reflective/Active values tend to be fairly different than those of the other members of the group. However given the specific learner space, all the assignments (except from learner 2) seem to be the best possible, as it is indicated by the membership probabilities. Regarding learner 2, he was initially assigned to the 4th Group but then the equalization function re-assigned her

Table 2: Test group formation using Random Selection Algorithm

Learner	Group	Sensing/ Intuiting	Reflective/ Active
1	1	-5	3
2	1	5	-9
3	1	11	1
4	2	7	1
5	2	11	-5
6	2	9	3
7	3	11	3
8	3	-11	1
9	3	7	3
10	4	-11	3
11	4	7	-9
12	4	11	-9
13	5	-5	1
14	5	-11	3
15	5	9	-9
16	6	-5	-11
17	6	9	3
18	6	7	3

to Group 1. In particular, the initial grouping process resulted with Group 4 having 5 members, i.e. with one more member than the maximum allowed members per group $n = round(18/5) = 4$. Then, the equalization function was activated to evenly distribute students to groups, i.e. the system had to re-assign one member of the 4th group. According to the equalization function described in Section 2.1, the best possible candidate was learner 2 who had the lowest probability (54.15%) of belonging to Group 4. As illustrated in Table 1, the next "best" group for learner 2 is Group 1 with membership probability 20.37%. Thus learner 2 is "moved" to Group 1 as the best possible re-assignment since the size of the group is less than the maximum allowed.

Heterogeneous grouping. In the second test we aim to assign the 18 learners to 6 heterogeneous groups based on two criteria; the Sequential/Global axis of the Felder & Silverman model and the Reflective/Active axis of the Honey & Mumford model. The resulted groups are presented in Table 2. The six groups present an adequate level of heterogeneity. Specifically in groups 3, 4 and 5 of Table 2 we can see that the highest possible difference exists within the members of each group for the "Sensing/Intuiting" criterion (values -11 and 11 or 9), whereas, all the groups formed constitute of learners with a mixture of negative and positive values on both criteria. Only one of the six groups, Group 2, presents a level of homogeneity based on one of the criteria, the "Sensing/Intuiting" criterion (since all the

learners' values are very close: 7, 11, 9). However, even this group presents also a satisfactory level of heterogeneity based on the "Reflective/Active" criterion.

5. Conclusions and Further Research

In this paper we presented the web-based Group Formation Tool which uses low complexity algorithms for homogeneous and heterogeneous groupings. Preliminary tests on grouping 18 learners showed that the level of homogeneity and heterogeneity of the resulted groups was quite satisfactory; however the evaluation of the tool in a real context is vital. Moreover the use of a wider range of criteria is imperative in order to draw safer conclusions about the adaptability and the robustness of the algorithms (especially in the case of FCM).

In the future we aim to investigate the development of a low-complexity, easy-to-implement mixed type (homogeneous and heterogeneous) grouping algorithm. To this direction, optimization algorithms like the Hill Climbing Algorithm as well as the FCM algorithm will be further investigated. Another interesting direction is to investigate the incorporation of weighted criteria to the FCM algorithm. This possibly needs to re-design the algorithm in a way that would enable the use of a weight exponent within the distance calculation function. Finally, we plan to (a) incorporate this tool to an e-learning environment as a module for forming groups or peer searching using multiple criteria, and (b) extend the tool allowing learners to select grouping criteria, especially in cases of searching for peer support, as well as intervening in the grouping process. In particular, the group negotiation option will be extended so that in case of agreement between learner and instructor, the learner's group preference would also be taken into consideration in re-assigning learners to groups or in the initial grouping process. All the above directions are quite challenging and further research is necessary so as to prove the effectiveness and usefulness of such a group formation module.

Acknowledgment

References

[1] J. C. Bezdek. *Pattern Recognition with Objective Function Algorithms*. Plenum Press, New York, 1981.

[2] R. Cavanaugh, M. Ellis, R. Layton, and M. Ardis. Automating the process of assigning students to cooperative- learning teams. In *Proceedings of the 2004 American Society for Engineering Education Annual Conference & Exposition*, June 2004.

[3] C. Christodoulopoulos and K. Papapanikolaou. Investigating group formation using low complexity algorithms. In *Proceedings of the Personalisation in E-Learning Environments at Individual and Group Level Workshop, in 11th International Conference on User Modeling, Available at http://hermis.di.uoa.gr/PeLEIGL*, pages 57–60, June 2007.

[4] A. Gogoulou, G. Gouli, E. Boas, E. Liakou, and M. Grigoriadou. Forming homogeneous, heterogeneous and mixed groups of learners. In *Proceedings of the Personalisation in E-Learning Environments at Individual and Group Level Workshop, in 11th International Conference on User Modeling, Available at http://hermis.di.uoa.gr/PeLEIGL*, pages 33–40, June 2007.

[5] S. Graf and R. Bekele. Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization. In *Proceedings of the 8th International Conference on Intelligent Tutoring Systems (ITS 2006)*, pages 217–226. Springer Berlin / Heidelberg, June 2006.

[6] J. Greer, G. McCalla, J. Cooke, J. Collins, V. Kumar, A. Bishop, and J. Vassileva. The intelligent helpdesk: Supporting peer - help in a university course. In *Proceedings of the 4th Int Conference on Intelligent Tutoring Systems (ITS 1998)*, volume 1452 of *Lecture Notes in Computer Science*, pages 494–503. Springer - Verlag: Berlin, June 1998.

[7] M. Hishina, R. Okada, and K. Suzuki. Group formation for web-based collaborative learning with personality information. *International Journal on E-Learning*, 4(3):351–364, 2005.

[8] A. Inaba, T. Supnithi, M. Ikeda, R. Mizoguchi, and J. Toyoda. How can we form effective collaborative learning groups? In *Proceedings of the 5th International Conference on Intelligent Tutoring Systems (ITS 2000)*, volume 1839 of *Lecture Notes in Computer Science*, pages 282–291. Springer-Verlag London, June 2000.

[9] E. Martin and P. Paredes. Using learning styles for dynamic group formation in adaptive collaborative hypermedia systems. In *Proceedings of the 1st International Workshop on Adaptive Hypermedia and Collaborative Web-based Systems*, pages 188–198, July 2004.

[10] G. Serban and G. S. Moldovan. A comparison of clustering techniques in aspect mining. *Studia Universitatis Babes-Bolyai Informatica*, LI(1):69–78, 2006.

[11] M. Vaziriannis and M. Halkidi. *Knowledge Mining from Databases (in Greek)*. Typothito - Georgios Dardanos, Athens, 2003.

[12] D.-Y. Wang, S. S. Lin, and C.-T. Sun. Diana: A computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups. *Computers in Human Behavior*, (23):1997–2010, 2007.

[13] X. Y. Wang and J. M. Garibaldi. A comparison of fuzzy and non-fuzzy clustering techniques in cancer diagnosis. In *Proceedings of the 2nd International Conference on Computational Intelligence in Medicine and Healthcare*, June 2005.

[14] I. Wilkinson and I. Fung. Small-group composition and peer effects. *International Journal of Educational Research*, 37:425–447, 2002.