# A Group Formation Tool for e-Learning Environments

Christos E. Christodoulopoulos
Technology Education and Digital Systems
University of Piraeus, Greece
christos.c@ieee.org

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

## Abstract

*Group formation in e-learning environments is a challenging area. In this paper we present a web-based group formation tool that 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 algorithm for creating homogeneous groupings, which provides for each learner the probability that s/he belongs to different groups. This information is also provided to the instructor when he needs to manually exchange learners or intervene 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. In this context, we think that it is important in the process of forming groups to take into consideration learners' individual characteristics 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 environments has lead to the development of computer-based tools that support automatic group formation based on learners' characteristics. 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], surveys are available to learners in order to provide the system with their personality traits and the instructor can select multiple criteria using weights for each one. The specific system uses the Hill Climbing algorithm in order to provide both homogeneous and heterogeneous groupings. In the OmadoGenesis tool [4], the instructor can select the learners for whom to apply the grouping, he/she can select a grouping type (homogeneous or heterogeneous) for each of the criteria used (up to 3) and he/she can edit the created groups. However the instructor cannot perform directly manual grouping but by using the edit function he/she can manually intervene in the proposed grouping. This tool performs 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. A useful option provided in the tool of Cavanaugh et. al [2] and OmadoGenesis [4] to the instructor is the option to manually define or edit the groups. 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 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. More 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 more than adequately for either homogeneous, heterogeneous or mixed-type (i.e. homogeneous for one attribute and heterogeneous for another) group formations, their computational complexity (polynomial) and hard implementation make them inappropriate 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 our previous work [3] we compared the two best-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 made them good candidates for supporting homogeneous groupings. In the course of our study both algorithms in their standard forms were evaluated in a simulated environment and it was proven that certain features of the "fuzzy" nature of the FCM make it promising for group formation in learning environments. These features will be presented later on.

Regarding the issue of heterogeneous grouping, and especially about the definition of heterogeneity in learning groups, research has been very limited until now. 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). An expanded definition more suitable for our multi-valued student space could describe the heterogenous group as a group where all the possible values of the student space are present. As discussed in a following section, this approach lead to the use of a random algorithm in order to provide heterogeneous groups with satisfying results.

The learner space model is multi-dimensional; each dimension represents a 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 in order to facilitate the grouping process and the instructor in intervening to the grouping procedure 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.

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

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 student model is multi-dimensional; every dimension represents one different personality attribute.

The FCM algorithm, as it was implemented in our application, is structured as follows:

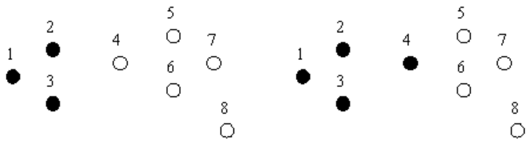
1. Initialize Probability Matrix $U$:
   (a) Assign random values to $U$
   (b) For every member $j$, make $u_j$ sums to 1
   $u_{ij} = u_{ij} / \sum_{j=1}^{n}(u_j)$
2. For every step $\kappa$:
   (a) Calculate new cluster centers for each cluster $i$:
   $$c_i = \frac{\sum_{j=1}^{n}(u_{ij}^{m} \cdot \vec{x_j})}{\sum_{j=1}^{n}(u_{ij}^{m})}$$
   (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

The main problem that we encountered during our tests with FCM and k-means algorithms was the inequality in the number of members between clusters. The source of this problem was 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 the others.

The challenge was to amend the situation using low complexity functions. To this end FCM's probability matrix proved extremely useful.

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. This solution enabled us to reduce the required amount of procedures to just basic conditional search. The equalization function is presented below:

1. Calculate the number of members per group:
   $$n = (round(total\_students/groups))$$
2. For every cluster $i$
   (a) Find if there are any extra members:
   if $count(c_i) > n$ then
   (b) For every extra member that has the lowest probability for this cluster:
   for $j$ where $u_{ij} = \min\{u_i\}$




(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

Intelligent and Adaptive Learning Environments

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

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

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 $x' \neq i$

An important "side effect" of this function is that by exchanging members between the groups based on the highest probability proposed by the FCM, we manage to result in groups that are equal in number with the best possible clustering quality, although it is lower than the initial. However the instructor is offered the option to activate this function and automatically redistribute students to groups or deactivate it and manually redistribute students to groups 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 idea behind this approach lies in the definition of a heterogeneous group. Although, as discussed above, there is no widely accepted definition, in our work we propose a strict definition of a heterogeneous group as the one that contains all the possible values of the learner space. Since such groups would require a large number of members (consider that in a 2-dimensional space with 5 possible values per dimension we would need $\simeq$ 7 members per group) and since it is not possible for the learners' attributes to cover the whole spectrum of the space, we pursued a more realistic approach. To ensure that every group could contain every possible value (but not all of them) we applied the previous definition to a probabilistic context. By using a uniform distribution we ensure that all the values of the search space have the same possibility to belong to a certain group. Therefore a random selection algorithm would provide groups that their members' attributes span across the search space. This algorithm provides an adequate level of heterogeneity while being extremely fast and easy to implement [3].

Note here that 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 examined against other heterogeneous group formation algorithms in real-life environments.

## 3. The Groups Formation Tool

The Groups Formation Tool was developed in PHP and uses a MySQL database. The database was built to support multiple lessons/instructors and the algorithms were implemented in a modular way in order to ensure the minimum possible dependence from the user environment. These features allow for the tool to be easily integrated with any e-class environment.

The instructor is able to form groups either manually or automatically. The instructor is asked to choose the desired grouping method (manual or automatic), and to en-

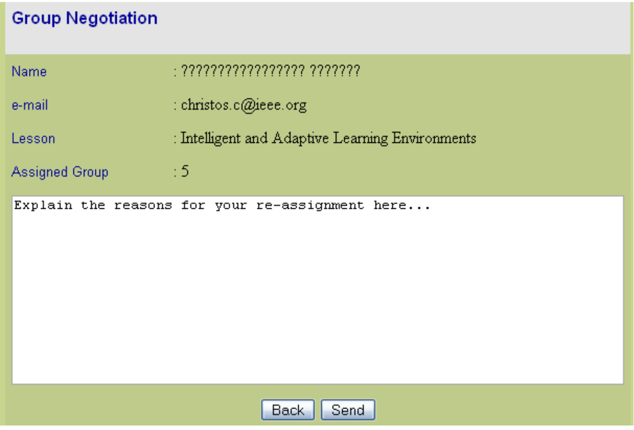



Figure 3: A screen shot of Groups Formation Tool (Learner Environment: Group Negotiation). The learner is allowed to negotiate his re-assignment to a different group through the 'Group Negotiation' form. The name of the learner has been deliberately replaced by question marks "?"

ter the number of groups he/she wants to create. Using the automated process, the instructor selects whether he/she wants to create homogeneous or heterogeneous groups and is asked to select up to three different criteria. Finally the instructor can choose whether the system is to equalize the number of the members in the created groups.

After the automatic group formation process, the resulted groups are presented to the instructor in the form of separate tables. Along with each learner's name the system provides a list of certain attribute values. The instructor can select the set of the personality attributes to appear, by a configuration menu. The system will then store his/hers choice and personalize all the group listings.

The instructor has always the option to edit the output of any grouping by using the Edit mode where he/she can re-assign the learners to different groups. In this mode, a more detailed profile for each learner appears along with the Membership Probabilities for each group, in case homogeneous grouping was selected (FCM algorithm). 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 is aware of a potential rivalry between two group members, or that an individual learner has asked to be re-assigned.

An example of the Edit mode is presented in Fig. 2. In this figure the members of the first group of the example shown in Table 1, are presented and for each member all the stored attributes are shown. In particular, the FCM generated membership probabilities, the knowledge level for the current lesson, the values for the Felder-Silverman Learning Style model, the Reflective/Active axis of the Honey-Mumford Learning Styles model and the Pragmatist/Theorist axis of the Honey-Mumford model. In this example the instructor wants to re-assign one learner from group 1 to another group. To achieve that, he/she must first pick the learner that has the least membership probability. In this case the second learner (with 20.37% membership probability) is the most appropriate candidate. Then the instructor finds the best possible group to assign the learner (in a similar way with the equalization function described above), according to the membership probabilities, which here is group no. 4 (with 54.15%). The instructor then simply replaces the group number in front of the learner and saves the new group. This method is by far more accurate and time-saving compared to the cross-checking of every attribute value of the learner.

The environment provides also a number of useful functions from a learner's point-of-view. A learner can view his/hers personality profile and alter it. In this case the instructor will be informed for any changes the next time he/she views the learner's profile.

A learner is also able to view his/hers group assignments in all the lessons that he/she participates, along with all his/hers group partners. Finally the learner can negotiate his/hers assignment with the instructor when he/she is not satisfied with it. Fig. 3 present this option. The learner must return a form to the instructor, stating the reasons for his re-assignment. The instructor can the re-assign the learner as he/she sees more appropriate.

## 4. Evaluation

In order to test 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. We must note at this point that the attribute values are far from being evenly distributed throughout the space. This "abnormality" adds an extra level of difficulty in the grouping process. This will be shown in the following two examples (one for homogeneous and another for heterogeneous grouping). Both examples were produced by the Group Formation Tool using the data from registered learners. The results were collected from the instructor interface of the Tool and are presented in tables 1 and 2. The first example test, presents a homogeneous group formation.

In this test the goal was to assign 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 six discrete values for each side of the axis:

Table 1: Test group formation using FCM and equalization function

| Learner | Group | Membership Probability | Sensing/ Intuiting | Reflective/ Active |
|---|---|---|---|---|
| 1 | 1 | 99.86% | -11 | 3 |
| 2 | 1<br>2<br>3<br>4<br>5 | 20.37%<br>7.34%<br>10.19%<br>54.15%<br>7.95% | -5 | -5 |
| 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 |

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 |

| | |
|---|---|
| 1(-1) | Very Low |
| 3(-3) | Low |
| 5(-5) | Moderate |
| 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 5 groups created, represent the most even distribution of the number of members per group, as 3 groups of 4 learners and 2 group of 3 learners were created. 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. That is because their Reflective/Active values tend to be fairly different than the ones of the rest of members in their group. However given the specific learner space, all the assignments (except the one of learner 2) seem to be the best possible, as it is indicated by the membership probabilities. Regarding learner 2, through the initial grouping procedure, was assigned to the 4th group and after the execution of the equalization function he was re-assigned to group 1. By the end of the initial grouping, group 4 had 5 members (one more than the maximum number of members per group $n = round(18/5) = 4$). In order to evenly distribute the number of members, the system had to re-assign one member of the fourth group. The best possible candidate was learner 2, who had the lowest probability (54.15%). As illustrated in Table 1, the next best group for learner 2 is 1 with 20.37%. Thus the system has made the best possible re-assignment.

Regarding the efficiency of the random selection algorithm, we present another example test. In this test we wanted to assign the same 18 learners into 6 groups using now the Sequential/Global axis of the Felder & Silverman model and the Reflective/Active axis of the Honey & Mumford model as our two criteria. The created groups are presented in Table 2. The groups present an adequate level of heterogeneity for both criteria. Specifically in groups 3,4 and 5 we can see that the highest possible difference exists within the same group for the "Sequential/Global" criterion (-11 and 11 or 9), whereas, all the groups formed constitute of learners with a mixture of negative and positive values on both criteria. One exception is the second group, where all the learners have values $\geq 7$ for the "Sequential/Global" criterion.

## 5. Conclusions and Further Research

In this paper we have presented a web-based group formation tool that satisfies the requirements for efficiency and low complexity, posed by the implementation of most e-learning environments. Our initial simulation tests, suggested that both the homogenous and heterogeneous algorithms have satisfying results; however the evaluation of this tool under real-life conditions is vital. Moreover the use of a wider range of criteria is imperative, as we will draw safer conclusions about the adaptability and the robustness of the algorithms (especially in the case of FCM).

One major challenge for our project will be the development of a low-complexity, easy-to-implement mixed-type (homogeneous and heterogeneous) grouping algorithm. The solution would have to derive most possibly from optimization algorithms by experimenting with modifications on the simpler ones like the Hill Climbing Algorithm. Experiments with the FCM algorithm with the appropriate use of the Membership Probabilities might also be promising.

Another interesting direction is to investigate the incorporation of weighted criteria to the FCM algorithm. That would require the re-design of 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 learner environment 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. This direction is quite challenging and further research is necessary so as to prove the effectiveness of such a module.

## Acknowledgment

The authors would like to thank Mr. Chrysovalantis Kefalidis, Mr. Christos Kyriakopoulos and Mr. John Petropoulos for their help with the implementation of the first version of the tool.

## 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 (to appear)*, 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 (to appear)*, 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)*, 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)*, 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.