Analyzing collaboration and interaction in learning environments to form learner groups

1. Introduction

In recent years there has been a huge increase in the use of systems that support collective learning processes in which groups of students collaborate to achieve common goals (Gress, Fior, Hadwin, & Winne, 2010). The research field of CSCL (Computer-Supported Collaborative Learnings) studies how to take advantage of technology to improve these collective processes (Dillenbourg, Jarvela, & Fischer, 2009). Thus, nowadays learners are used to holding discussions within the subjects they are studying in forums and sharing tasks and materials with other students. CSCL systems can also enable collaborative problem-solving activities (Alavi & Dillenbourg, 2012) by providing shared workspaces where learners create solutions (source codes of computer programs, models, etc.) that solve specific tasks proposed by the teacher.

This proliferation of collaborative systems has been made possible thanks to the understanding of how these collective processes are conducted (Kahrimanis, Avouris, & Komis, 2011). Bravo, Redondo, Verdejo and Ortega (2008) follow the following three phases for analysing collaboration among learners: observation, abstraction and intervention. In the initial observation phase, raw data about the interaction of each learner with the collaborative system are stored (i.e., proposals of each learner, agreements or disagreements in the group, kind of actions performed, manipulation of shared artefacts, etc.). In the abstraction phase, a set of variables, known as analysis indicators (Anaya & Boticario, 2009), are calculated to assess different aspects of the collaboration (e.g., coordination between students, level of communication, quality of the solution built by the learners, etc.). Finally, the intervention phase (Duque, Bravo, & Ortega, 2013) uses the analysis indicators to design actions that improve the collective process. These interventions can include, for example, advice to learners, changing the difficulty of the task or opening a discussion about a topic.

A specific type of intervention is the use of analysis indicators from previous activities to form new learner groups, for example by grouping learners of a similar productivity on previous tasks, or by putting learners of a similar level of knowledge together or by ensuring every group has at least one student with very high marks, etc. The complexity of designing work groups can be high, as many analysis indicators may need to be evaluated and the number of possible groups grows exponentially as the number of learners increases. Moreover, different approaches should be considered for selecting the learners in each group. For instance, the teacher could require heterogeneous group in which the learners have different skills, or the teacher may prefer homogeneous groups in which the learners of a group have the same abilities or knowledge level to carry out productive collaboration. Several studies (Samsudin, 2006) have been carried out to evaluate heterogeneous and homogeneous grouping in collaborative learning with disparity of conclusions about their strengths and weaknesses. This divergence of conclusions motivates the construction of a flexible method that would enable teachers to choose the criteria that best fit a specific learning setting, and that would also permit them to combine homogeneous and heterogeneous criteria in a mix approach that consists of groups in which the learners have a common set of skills (fluent communication, work speed, etc.) but they are different in other aspects (accuracy of the solutions, quality of the documentation, etc.). At this point the concept of data depth (Zuo & Serfling, 2000) is used to measure how close the analysis indicators' values are for a learner compared to the values that the same indicators take for the other learners. The data depth is a measure of how central a given datum is with respect to a distribution function or a given dataset. Thus, the data depth organizes a given dataset in the following way: if a datum is moved toward the centre of the data cloud, then its depth increases, and if the datum is moved toward the outside, then its depth decreases. This paper proposes to take advantage of the data depth concept to design a flexible method that forms learner groups taking into account different analysis indicators and that can be adapted to the requirements of creating homogeneous or/and heterogeneous groups. A software tool has also been built to automatically perform these processes of forming learner groups. The tool performs a data-driven decision making process in which data about the collective processes are used to decide which are the most suitable groups to be formed to approach future tasks would be. Therefore, this proposal is not only a theoretical one but it also can be applied in learning settings with a large number of students and analysis indicators. For this purpose, this paper describes two case studies that have been carried out to take advantage of the tool and to put the proposal into action. These case studies focus on group formation of learners who solve academic tasks in different domains (computer programming and data mining).

This paper is organized into five sections. Section 2 reviews the related work aimed at forming learner groups in collaborative systems. Section 3 describes our method based on the concept of data depth for grouping learners in different domains (computer programming and data mining). Section 4 presents two case studies in which our method is applied to form learner groups for solving programming and data-mining tasks. Finally, Section 5 discusses the conclusions drawn from the work.

2. Theoretical Background

The CSCL research area is focused on how collaborative learning supported by technology can enhance peer interaction and group work (Lipponen, 2002). CSCL systems enable the learners to interact amongst themselves, without the constraint of time and space as would be the case in a real classroom situation (Ho, Shyu, Wang, & Li, 2009). In this context, the challenge of forming learner groups whose members achieve productive collaboration arises (Kreijns, Kirschner, & Vermeulen, 2013).

One of the criteria used to form groups of learners is to consider the level of the students. Thus, Mathews (1992) points out that students with high-level abilities prefer homogeneous groups. This is situation was explained by Abrami, Chambers, Poulsen, De Simone, d'Apollonia, and Howden (1995) who indicate that homogeneous groups of students are beneficial because good students are able to perform their activities better if they do not have to spend time on explaining tasks or answering questions to other students. In a similar way, Baer (2003) concludes that the academic performances of homogeneous learner groups are better than those obtained by heterogeneous groups. However, Wang, Lin, and Sun (2007) reviewed an important number of research works and concluded that groups with higher levels of intra-group diversity provide other benefits such as greater student collaboration in task-solving, and that classroom distribution is not polarized into groups of good students and lower-level students who are more hesitant to interact with others (Webb, 1985).

The characteristics of the tasks to be solved can be used as other criteria to form the learner groups. Thus, Paredes, Ortigosa, and Rodriguez (2009) conclude that projects that include a wide range of tasks are especially suitable for heterogeneous groups, while homogeneous groups might be better at achieving specific goals. Moreover, West (2002) points out that heterogeneous groups generate different point of views about how to solve a problem and this diversity can be particularly useful in creative tasks.

Personal attitudes, learning styles and the preferences of each student are also factors to consider when establishing these working groups. According to Johnson and Johnson (1989), heterogeneous groups improve the acceptance of other cultures and behaviours. Faris (2009) has focused his analysis on comparing the academic performance of learner groups with different skills, nationalities and cultural backgrounds in comparison to homogeneous learner groups with similar skills and cultural backgrounds. This author concludes that the heterogeneous learners should dialogue and negotiate among themselves to come to an agreement and this is a contribution to the academic formation process. The same conclusions are achieved by Wang et al. (2007) who used the proposal of Sternberg (1994) to group students heterogeneously in order to achieve better results than a randomly formed group. The number of psychological variables and the type of learning styles can generate multiple forms of student groups. Moreover, the attitudes of the learners can change during the task development. Thus, Worchel, Rothgerber, and Day (2009) indicate that the phenomenon of social loafing in which a learner has the tendency to reduce his/her effort is more common in later stages of the tasks. For those reasons, it is necessary to build computational support that carefully analyses the learners' behaviour and performs the group formation according to a high number of factors.

Ho et al. (2009) have proposed an algorithm that forms highly heterogeneous groups using the criterion of learning style, competence and student interaction. The results of this algorithm have been analysed from a computational point of view (time, implementation, etc.). However, this analysis does not include a study of the pedagogical consequences of forming extremely heterogeneous groups. Christodoulopoulos and Papanikolaou (2007) have studied the application of clustering algorithms, which are designed to discover groups of data, to form learner groups. This study is only focused on identifying the most suitable algorithm, and it does not take into account the tools that can apply to different approaches to group formation according to the preferences of the teacher and the learning setting. Therefore, this paper approaches the challenges of proposing a method that allows teachers to define different criteria in order to form groups using the analysis indicators generated in previous tasks.

Group formation can also be a process in which the learners can participate by showing their preference for certain partners. In order to form learner groups automatically, the iHUCOFS (Integrated Human Coalition Formation and Scaffolding) module (Khandaker & Soh, 2010) uses an analysis of the capabilities and preferences of each student. Muehlenbrock (2005) developed a module that analyses learner profiles and contexts to immediately match learners who can work together. These approaches are focused on forming groups, but they do not use the previous results of the analysis of the collective work.

3. Using Tukey Data Depth to form groups of learners

The concept of statistical depth has received a lot of attention from statistical researchers during the last decade, as is seen in the book by (Liu, Serfling, & Souvaine, 2008). However, it has not yet been applied in most fields. In fact, to the best of our knowledge, it is the first case in which data depth is applied in the formation of learner groups in collaborative learning environments. The importance of this statistical concept arises from the fact that it generalizes the one-dimensional idea of mean or median to higher dimensional spaces.

When using only one analysis indicator (one-dimensional space), two definitions of data depth exist: the Tukey depth (Tukey, 1975) and the simplicial depth (Liu, 1990). Let X be the analysis indicator value taken by a given learner, and let X_1, ..., X_n be the analysis indicator value taken by the set of learners with respect to whom the depth is computed. Then, the Tukey depth of X with respect to X_1, ..., X_n is two over *n* times the minimum between the amount of values, X_1, ..., X_n, smaller than or equal to X and the amount of values larger than or equal to X. The simplicial depth of X with respect to X_1, ..., X_n is two over *n* times the product of the amount of values, X_1, ..., X_n, smaller than or equal to X and the amount of values larger than or equal to X. From these definitions it is easily confirmed that in a one-dimensional space, the deepest point coincides with the median, and also for instance the mean when following a normal distribution.

Although the mean or median is sufficient to understand this idea of centrality when we are only considering one analysis indicator, see the left part of Figure 1, they are not problem-free for several indicator at the same time. It is well known that the median calculated for each analysis indicator may not always give an idea of depth, and in some cases it may not even fall inside the cloud formed by the learners. As an example, in the right-hand side plot of Figure 1, we have four learners represented by two analysis indicators which are the vertices of an equilateral triangle and its center of mass. By symmetry, the deepest point should be unique and coincide with the center of mass. However, that is not the case of the median points, represented by the green zone in the plot. Moreover, the depth is affine invariance, but this is not the case of the, coordinate-wise, median. It is easy to see that this zone varies if we rotate the axis, i.e., if in the right-hand side plot of Figure 1, we fix the four, square red, points but rotate the, blue, coordinate axes, the median will be now computed with respect to the new axes, given a different, green, zone of median points. The problem particularly is aggravated if we go further from multidimensional spaces to functional spaces, i.e. having curves, which depend of the time for instance, instead of multidimensional points. That is, when we consider analysis indicators that depend on time, it is aggravated due to the coordinate-wise, mean or median of a set of curves neither necessarily has the shape of the curves in the dataset nor fulfil its properties. However, this is out of the scope of this paper as here we concentrate ourselves in multidimensional spaces.

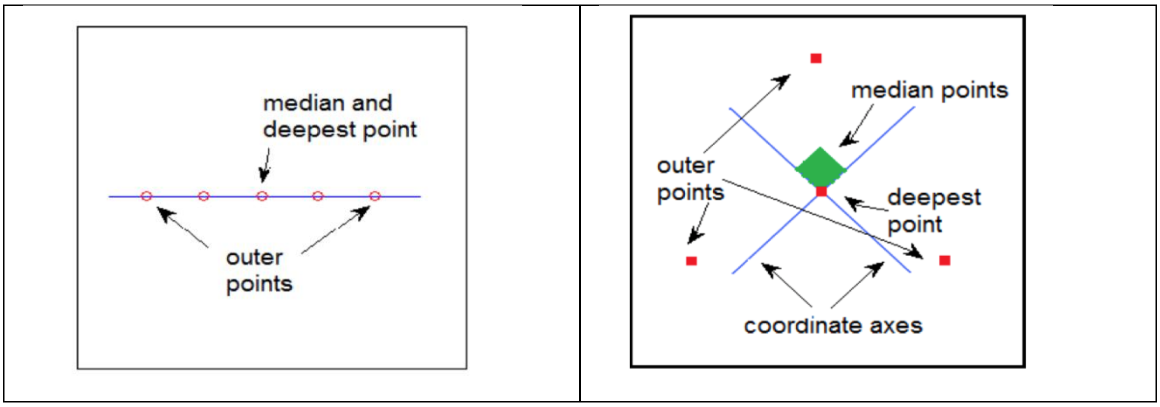

Figure 1. Graphical representation of the data depth concept.

Taking that into account, to measure centrality we need to use the concept of depth, which for multidimensional spaces, was formalized by Zuo and Serfling (2000). There, some requirements that every depth should fulfil were established. Some of the best-known proposed depths are the Tukey depth, and its computable version, the random Tukey depth (Cuesta & Nieto-Reyes, 2008), Oja's depth (Oja, 1983), and the simplicial depth.. The reason for using the random Tukey depth is that it is amongst the best-known definitions of depths, and the Tukey depth is the only one with a computationally effective version in any type of dimension as it is pointed out by Cuesta and Nieto-Reyes (2008). In practice, the above cited depths can only be computed in low dimensions, such as dimensions two and three.

Thus, given a dataset of learners, in which each learner has assigned some analysis indicators values, the data depth of a learner is a quantity that measures how central the learner is. The data depth of a learner increases when it is closer to the "center (of the dataset)". Notice that by this definition, data depth identifies outliers, i.e. learners that do not fit into the classroom pattern. In this article, we propose using data depth to organize a given dataset as the first step of forming groups. The "deepest" learners will be the last to be grouped due to their lack of special needs. Notice that the "deepest" learner is also the most representative individual in the dataset. Another way of understanding this problem is to identify each learner with a vector and then to find the point that is more inside the point's cloud or roughly speaking, the one with more points around it.

The data depth is derived from a probability measurement and takes values in the interval [0,1]. However, when the problem at hand allows for it, it is common to work with the ranks associated with this measurement instead. When ties cannot occur, a random tie breaking procedure is generally applied. For instance, an analysis indicator assesses the level of *communication* of each learner in previous tasks with a numerical value between 1 (the lowest level of communication) and 5 (the highest level of communication). There are eleven students and the *communication* indicator takes the values collected in Table 1. These data have been ordered to identify the rank associated to their data depth. Thus, the value 3.5 is the deepest due to it being the central value of the set of data while the values 1 and 4.9 are the least deep as they are the most external. The depth rank of each value is evaluated by means of a range of positive integer numbers in which the highest integer is assigned to the central value and the lowest integer is associated to the most external value(s). The data depth enables us to apply a criterion by identifying which learners have similar data depth. When applying a concentration criterion, the learners in a group should have similar data depth and analysis indicators, whereas when applying a dispersion criterion the learners should have different analysis indicator values while still having similar data depth. Taking into account that the concentration criterion can be applied to a set of analysis indicators and the dispersion criterion with respect to another set, both criteria can be performed simultaneously. Note that in this case the data depth is computed two times: for applying the concentration criterion and then for the dispersion criterion.

Table 1. An example of the data depth of the analysis indicators registered by learners.

Learner identification	Value of the communication indicator	Data depth ranks
11	1.0	1
9	1.4	2
6	2.0	3
5	2.7	4
3	2.9	5
1	3.5	6
2	3.9	5
4	4.1	4
7	4.3	3
10	4.8	2
8	4.9	1

Our method makes use of the concept of data depth to arrange the learners in a certain way, which will help to group them according to the criteria given by the teacher. The idea is to select the learner groups iteratively, adding one learner to a group per round in such a way that a certain value, which is the score given to an arrangement of learners in groups, is minimized. The formula is as follows, $Score(A) = \sum_{G \in A}(V_1\,(G) + E_2(G)^2)$, where $V_1(G)$ is the variance of the group considering only the indicators that should be homogenous and $E_2(G)$ is the mean of the analysis indicators that should be heterogeneous.

The reason for choosing this score function is that when the group follows the criteria of the teacher, the variance in the homogenous analysis indicators and the mean of the heterogeneous analysis indicators are smaller in the new groups. The idea behind using this formula is that the score function can be simplified as a function depending exclusively on the variance of the homogenous analysis indicators and the variance of the heterogeneous analysis indicators in the groups of the arrangement.

The proposed algorithm implements the following steps:

1. Calculate the data depth of the learner dataset with respect to homogenous analysis indicators and heterogeneous analysis indicators. Treat this pair as a vector.
2. Organize the learners using the pair of values of data depth with respect to homogenous analysis indicators and heterogeneous analysis indicators.
3. Generate an arrangement with empty groups.
4. For each learner:
    a. Add the learner to the group that minimizes the score function.
5. Return the arrangement.

Note that the output is heavily dependent on the way the learners are put into order in the second step. There is not a canonical way to sorting vectors, we propose using a clustering algorithm. The clustering algorithm will put learners with similar data depth in both sets of indicators together and then, selecting the group with smaller norm vectors, we sort out their learners in any order. This process continues until all the learners are classified.

4. The method in action

After designing the method used to form groups of learners in collaborative environments, we carried out two studies to put this proposal into action. The analysis indicators used in the first study have been proposed by Bravo et al. (2008) who carried out a process to select suitable indicators. The teachers of the learners had an active role in the process of selection of indicators, and they indicated that the indicators were suitable to analyze the learner as an individual, as well as the group and the solution to the tasks. In the second study, the analysis indicators take values according to the subjective point of view of each learner. For instance, a certain amount of time solving a task can be considered as very high by one learner and as very low by another learner. However, this second study shows that the satisfaction of the learner will be also better when the analysis indicators take a value assigned by the students.

A software tool was developed to automatically perform the process of group formation in these studies, which are described in the following subsections. According to Wang et al. (2007), heterogeneous groups are better at performing the tasks, although these groups should not be overly heterogeneous. To achieve this goal, the teachers define a set of analysis indicators which are heterogeneous but they also define analysis indicators whose values are homogeneous in order to generate moderately heterogeneous groups.

4.1. Collaborative programming

The proposed method was used in this first case study to form groups of learners that had to solve programming tasks. There were one hundred and one learners, students of the Computer Science degree at the University of Castilla-La Mancha. They had previously solved several programming task using the COLLECE system (Bravo, Duque, & Gallardo, 2013), which allows learners to work collaboratively. This collaboration is distributed among learners. In other words, each member of the group works on his/her own computer and COLLECE provides a shared editor (Figure 2-a), coordination tools to edit (Figure 2-b), compile (Figure 2-c) and run (Figure 2-d) the programs on the console (Figure 2-e) and a chat feature (Figure 2-f) to allow communication between learners.

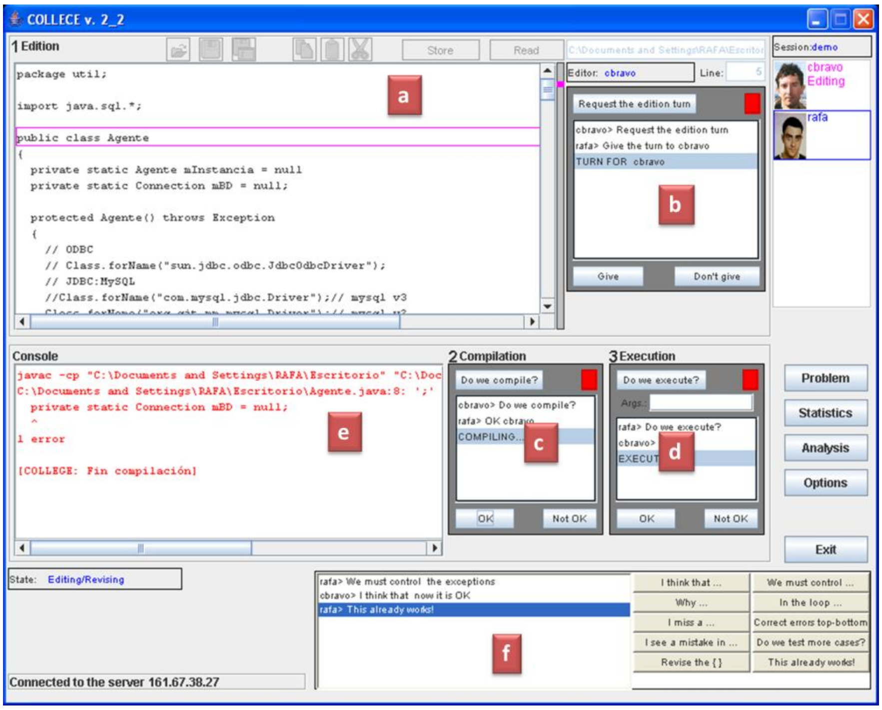

Figure 2. User interface of COLLECE.

COLLECE also includes an analysis subsystem that infers a set of analysis indicators which assess the collective work. For this purpose, the analysis subsystem collects the actions performed by the learners and the programs they build. The subsystem then applies mechanisms based on fuzzy logic to infer indicators that analyse the learners' activity. These mechanisms use a set of rules that were designed taking into account the proposals of teachers of Programming subjects at the University of Castilla-La Mancha about how to evaluate these learning activities (Bravo et al., 2008). Each indicator is assigned a value in the set of integers between 1 (the lowest value) and 5 (the highest value). The analysis indicators can be classified in the following three categories according to their purpose:

- Analysis indicators of the learner: They assess the activity of each member of the group. Examples of indicators that fall into this category are *work* and *discussion*. The *work* indicator evaluates the dedication of the learner in solving the task. The *discussion* indicator assesses the level of participation in the debate and the amount of exchange of ideas with the rest of the group.
- Analysis indicators of the group: They evaluate the behaviour of each group. For example, two of these indicators are *coordination* and *speed*. The coordination

indicator measures the degree to which the members of the group usually agree on how to share the workload and workspaces. The *speed* indicator assesses the amount of time spent on a task.

- Analysis indicators of the solution: These evaluate the program built to solve the task. For example, some of these indicators are *well-formed* and *quality*. The *well-formed* indicator evaluates if the program is syntactically correct. The *quality* indicator assesses whether the program generates the suitable output.

After carrying out several programming tasks with COLLECE, in which learner groups were randomly formed, the proposed method was used by the teacher to form new groups. The teacher decided to generate 34 groups, in which the members within each group should have similar values (concentration criterion) for the *work* and *discussion* indicators that evaluate individual activity; whereas the values for the *well-formed* and *quality* indicators, which evaluate the solutions they produced should be different (dispersion criterion). This results in groups formed by learners who have similar approaches to their work while they usually achieve different results. Thus, the teacher encourages the learners to maintain a discussion and explain their different views while creating the programs. Previous studies (Bravo et al., 2013) have been carried out to analyse the advantages and/or disadvantages of applying collaborative programming practices (for example, pair programming practices) to the solo programming practices of learners. However, these studies did not take into account a method that forms groups of programmers according to criteria that benefit the learning process. For this reason, we outlined a case study in which the proposed method is used to support and complement the teacher's preference in forming groups.

The software tool that performs the process of group formation required the user to introduce as input: the number of groups to be formed, a CSV file with the values taken by a set of analysis indicators that assess different aspects of the collective learning processes previously carried out and the criterion (concentration or dispersion) to apply to each indicator. This file is structured in the form of a matrix, in which each column represents an analysis indicator and each row is made up of the values that the analysis indicator assigned to a particular learner in a specific task. This tool produces three pieces of output: a list, a graph and a table. The first piece of output is a list of the learners and the group which they should be included in. The second piece of output (Figure 3) is a graphical representation of the data depth rank of each learner. An identification number represents the learners in a two-dimensional Cartesian coordinate system. The x-axis represents the data depth rank of the learner with respect to the indicators that required a dispersion criterion. The y-axis represents the data depth rank of the learner respect to the indicators that required a concentration criterion. This graph is useful in interpreting the results. Note again that when the depth of a student is small, this means that the student is far from the "median" student. As an example, observing the horizontal line that goes through the coordinate (0,0.5), we see that the learners 50 and 49 are homogeneous with respect to the indicators of the y-axis but that they disperse with respect to the x-axis. This implies that both learners behave as "typical" students in the homogenous indicators, but this change when looking at the other sets of analysis indicators. This graphical aid can give an idea of who progresses at a different pace than the rest of the classroom. This is inspired by the DD-plot (Liu, Parelius, & Singh, 1999).

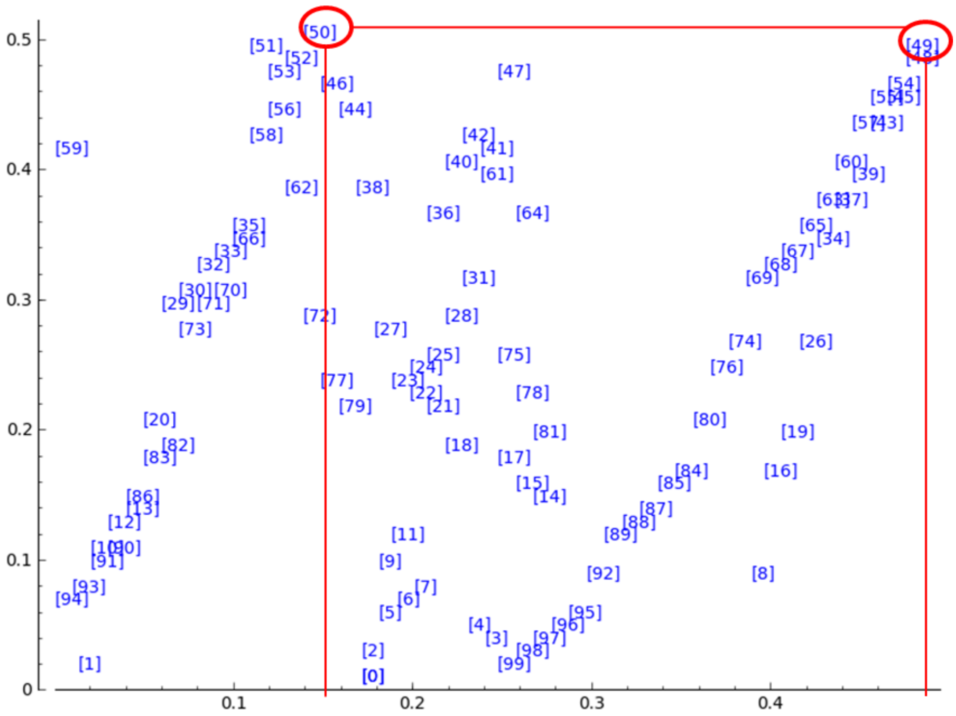

Figure 3.Graphical representation of the learners' data depth ranks.

The third piece of output of the tool is a table that includes the mean (M) values that the indicators take in each group and its standard deviation (SD). This statistical information provides a set of measurements about the degree to which the concentration and dispersion criteria are achieved. Table 2 shows an excerpt of the results provided by the tool in this case (for clarity purposes, the table includes the data of only the first 5 of the 34 groups). The value of the SD is especially important to give us an idea of the degree of dispersion and concentration of the analysis indicators' values in the new groups. For this reason, the last row of the table includes the average value of the SD for all groups (this average value takes into account the 34 groups, not just the 5 groups included in Table 2). In this way the table provides feedback to the teachers to be able to the degree to which their requirements have been satisfied. For example, in this case the SD of the *well-formed* and *quality* indicators are higher than the SD of the *work* and *participation* indicators. This can be considered a confirmation that the values of the *quality* and *well-formed* indicators fulfil a dispersion criterion, while the members of each group have similar values in *work* and *participation* (concentration criterion).

Table 2. Values of the analysis indicators in the new groups.

Group	Participation		Work		Well-formed		Quality	
	SD	M	SD	M	SD	M	SD	M
1	0,00	3,00	0,00	4,00	0,00	5,00	0,00	4,00
2	0,00	3,00	0,00	4,00	0,00	5,00	0,00	4,00
3	0,50	2,50	1,00	3,00	2,00	3,00	1,00	3,00
4	0,00	3,00	0,00	4,00	0,00	5,00	0,00	4,00
5	0,00	2,00	0,00	2,00	0,50	2,50	0,00	2,00
….								
Average values	0,12	2,70	0,25	3,00	0,53	4,31	0,42	4,33

4.2. Collaborative building of data-mining scripts

After carrying out a case study that evaluated the proposed method as a tool to fulfil the teacher's requirements to make groups of students combining criteria of homogeneity and heterogeneity, we designed a study to be able to assess the evolution of the learners' analysis indicators when the proposal was applied. For this purpose, twelve students from fourth and fifth course of the Computer Science degree at the University of Cantabria participated in a new case study. These students had to solve two tasks aimed at building scripts that had to perform data-mining processes. The first task was carried out by groups of two learners which were randomly formed. Learner groups worked in an asynchronous way. They divided the tasks into several subtasks which were assigned to one member of the group. Then, they used communication tools (e-mail, chats, etc.) to interchange ideas and questions and combine their results.

After performing this task, each learner completed a questionnaire to evaluate the collaborative worked carried out. This questionnaire asked the following question:

- Time: Amount of time spent on solving the tasks.
- Communication: Number of proposals made to the partner.
- Test: Number of testing processes carried out to verify the script.
- Participation: Number of initiatives to change features built by their partner.
- Work: Subjective evaluation of the amount of work carried out.
- Satisfaction: The subjective degree of satisfaction with their partner's work.
- Initiative: This measures the number of times in which the student has proposed solutions to the difficulties with the assignment.

For the experiment, the teachers decided that the students would be grouped under the following criteria: the indicators of *communication* and *initiative* in the same group to be as homogenous as possible; whereas the other indicators would be as heterogeneous as possible. The teachers did it in this way, because they felt that groups which there is a more active participant tend to show a clear asymmetry in the distribution of their tasks, i.e. the "leader" makes most of the decisions without consulting his or her partner. After performing the first task, the teachers asked the students to complete the same questionnaire again. Table 3 shows a comparison of the average values taken by the indicators in each task. The indicators that evaluate the collaboration between the members of the group (communication, initiative and participation) have higher values in the second task. The indicators of the amount of work, time and testing of the scripts have lower values in the second tasks.

Table 3. Average values of the analysis indicators in the tasks.

Analysis indicator	Average value in task 1	Average value in task 2
Time	2.583	1.000
Communication	2.500	2.750
Test	1.667	2.000
Participation	1.250	3.167
Work	2.000	1.667
Satisfaction	1.250	3.167
Initiative	1.333	2.750

Figure 4 shows that all new groups are more satisfied in the second task (this group was formed using the proposed method) than in the first one (when the group was formed randomly). We have used the Wilcoxon rank sum test to test the null hypothesis of the satisfaction of the groups. In the initial arrangement values are higher than or equal to the depth-based arrangement in contrast to the alternative hypothesis of satisfaction being higher when the groups are formed using our procedure. The p-value obtained is less than 0.001 so the null hypothesis is rejected. Therefore, we can conclude that the method enabled the formation of groups of learners who collaborate more fluently and need to employ less effort to solve the task. Moreover, their satisfaction is higher with the new group formed in the second task.

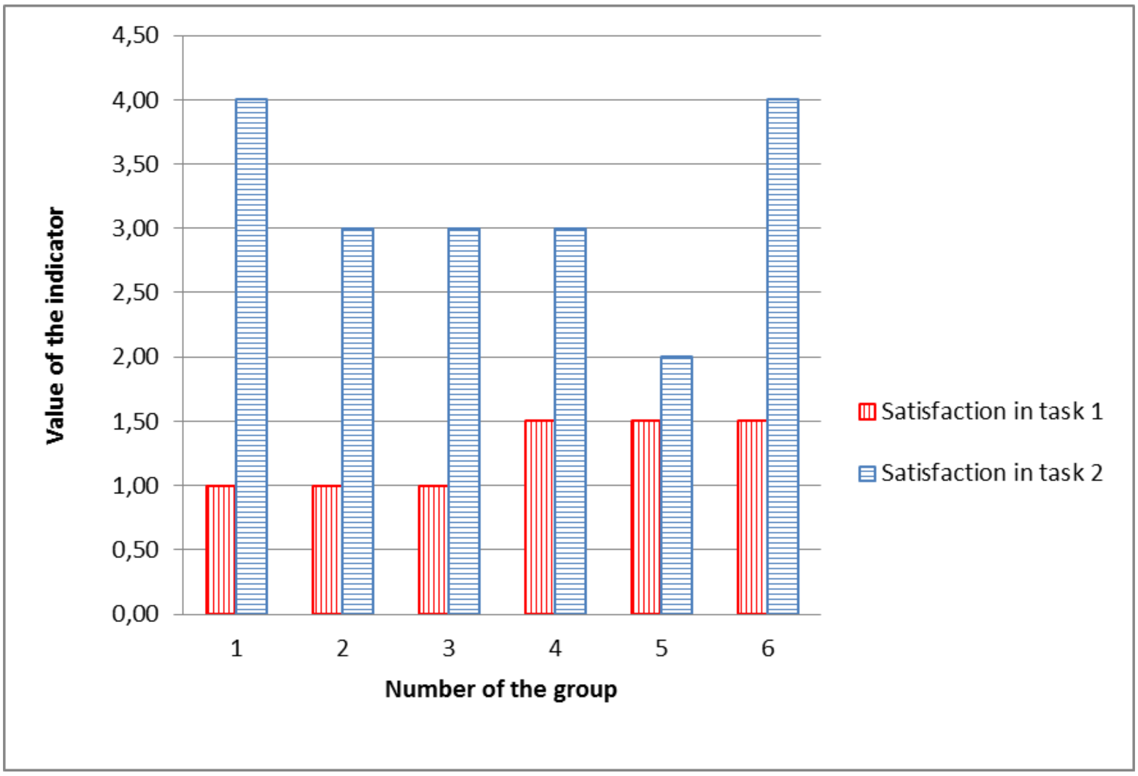

Figure 4. Learner satisfaction in the tasks.

5. Conclusions

This paper has presented a method to form learner groups that collaborate using CSCL systems. This method was applied and tested in two different contexts to form learner groups that had to solve tasks in the field of computer programming and data mining. The method uses indicators that analyse the activity previously carried out by the learners. The user specifies if the learners of the same group have to have similar or different values for certain indicators.

The proposal offers the advantage of combining two different approaches in the group formation process. Firstly, the learners of the group can be heterogeneous with respect to some skills or attitudes where they have different values in certain indicators. Secondly, a homogeneous criterion can simultaneously be applied in which learners registered similar values in other indicators. A software tool takes a file with the values of the analysis indicators available as input, and then automatically enacts this method. Then, the user specifies the number of groups to be formed and the tool produces a proposal of the members to be included in each group. Moreover, this tool shows a graphic representation of the data depth ranks of each learner and a statistical study of the values assigned to the analysis indicators.

Two case studies have been carried out to put our proposal into action. The first case study applies the method for forming group of learners in a Computer Science degree in which the students had to collaborate to solve programming tasks. This case study allowed the teacher to specify their own preferences in a flexible way in order to take advantage of the values of previous analysis indicators to form new learner groups. To this end, the software tool enabled the teacher to perform the process of group formation by applying concentration and dispersion criteria to the analysis indicators. The second case study focused on a comparison of the values assigned to the analysis indicators when the learners collaboratively solve a task to build a data-mining script, and when they are grouped by means of the proposed method. The learners were more satisfied with the groups formed by the proposed method and their collaboration was more productive.

According to these studies, the tool offers flexible support that enables teachers to define the criterion of concentration and/or dispersion to be applied to each analysis indicator. The two studies generated promising results with heterogeneous groups controlled by the teachers according to their own preferences and according to the satisfaction of the learners. In order to reduce the effort of teachers in deciding how to apply the criterion of concentration and dispersion, new research has been planned to provide a methodology that would include a set of steps, not only to generate analysis indicators, but also to specify the most suitable criterion of heterogeneity or homogeneity to be used in each indicator.

References

Abrami, P.C., Chambers, B., Poulsen, C., De Simone, C., d'Apollonia, S., & Howden, J. (1995). *Classroom Connections: Understanding and Using Cooperative Learning*. Toronto: Harcourt Brace & Co.

Alavi, H. S., & Dillenbourg, P. (2012). An Ambient Awareness Tool for Supporting Supervised Collaborative Problem Solving. *IEEE Transactions on Learning Technologies*, 5 (3), 264-274.

Anaya, A. R., & Boticario, J. (2009). A Data Mining Approach to Reveal Representative Collaboration Indicators in Open Collaboration Frameworks. *Proceedings of the 2nd International Conference on Educational Data Mining* (pp 210-219). Cordoba, Spain.

Baer, J. (2003). Grouping and achievement in cooperative learning. *College Teaching*, 51(4), 169-174.

Bravo, C., Duque, R., & Gallardo, J. (2013). A groupware system to support collaborative programming: Design and experiences. *Journal of Systems and Software*, 86(7), 1759-1771.

Bravo, C., Redondo, M. A., Verdejo, M. F., & Ortega, M. (2008). A framework for process-solution analysis in collaborative learning environments. *International Journal of Human-Computer Studies*, 66(11), 812-832.

Christodoulopoulos, C. E., & Papanikolaou, K. (2007). Investigation of group formation using low complexity algorithms. *Proceedings of the workshop on Personalisation in e-Learning Environments at Individual and Group Level at the 11th International Conference on User Modeling*, (pp 57–60). Corfu, Greece.

Cuesta-Albertos, J. A., & Nieto-Reyes, A. (2008). The random Tukey depth. *Computational Statistics & Data Analysis*, 52(11), 4979-4988.

Dillenbourg, P., Jarvela, S., & Fischer, F. (2009). The evolution of research on computers-supported collaborative learning: from design to orchestration. *Technology-Enhanced Learning*, *Springer Science+Business Media*, 3-19.

Duque, R., Bravo, C., & Ortega, M. (2013). An ontological approach to automating collaboration and interaction analysis in groupware systems. *Knowledge-Based Systems*, 37, 211-229.

Faris, A. O. (2009). The Impact of Homogeneous vs. Heterogeneous Collaborative Learning Groups in Multicultural Classes on the Achievement and Attitudes of Nine Graders Towards Learning Science. *Online Submission. Retrieved from ERIC database* (ED504109).

Gress, C. L. Z., Fior, M., Hadwin, A. F. , & Winne, P. H. (2010). Measurement and assessment in computer-supported collaborative learning. *Computers in Human Behavior*, 26(5), 806-814.

Ho, T. F., Shyu, S. J., Wang, F. H., & Li, C. J. (2009). Composing high-heterogeneous and high-interaction groups in collaborative learning with particle swarm optimization. In *Proceedings of the 2009 WRI World Congress on Computer Science and Information Engineering* (vol. 4, pp. 607-611). IEEE Computer Society.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Interaction Book Company.

Kahrimanis, G., Avouris, N., & Komis, V. (2011). Interaction analysis as a tool for supporting collaboration: An overview. *Technology-Enhanced Systems and Tools for Collaborative Learning Scaffolding Studies in Computational Intelligence*, 350, 93-114.

Khandaker, N., & Soh. L.-K. (2010). Improving Group Selection and Assessment in an Asynchronous Collaborative Writing Application. *International Journal of Artificial Intelligence in Education*, 20(3), 231-268.

Kreijns, K., Kirschner, P. A., & Vermeulen, M. (2013). Social aspects of CSCL environments: A research framework. *Educational Psychologist*, 48(4), 229-242.

Lipponen, L. (2002). Exploring foundations for computer-supported collaborative learning. In *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community* (pp. 72-81). International Society of the Learning Sciences.

Liu, R. Y. (1990). On a notion of data depth based on random simplices. *The Annals of Statistics*, *18*(1), 405-414.

Liu, R. Y., Parelius, J. M., & Singh, K. (1999). Multivariate analysis by data depth: descriptive statistics, graphics and inference. *The annals of statistics*, 27, 783–858.

Liu, R., Serfling, R., & Souvaine, D. L. (2008). Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications. *Dimacs Series in Discrete Mathematics and Theoretical Computer Science*, 72.

Matthews, M. (1992). Gifted students talk about cooperative learning, *Educational Leadership*. 50 (2)

Muehlenbrock, M. (2005). Formation of Learning Groups by using Learner Profiles and Context Information. *Proceedings of the 2005 conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology, IOS Press*, (pp 507-514). Amsterdam, The Netherlands.

Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statistics & Probability Letters*, *1*(6), 327-332.

Paredes, P., Ortigosa, A., & Rodriguez, P. (2009). 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)*.

Samsudin, S. (2006). Cooperative Learning: Heterogeneous Vs Homogeneous Grouping. In *APERA Conference*, *Hong Kong*.

Sternberg, R. J. (1994). *Thinking styles: Theory and assessment at the interface between intelligence and personality*.

Tukey, J. W. (1975, August). Mathematics and the picturing of data. In *Proceedings of the international congress of mathematicians* (Vol. 2, pp. 523-531).

Wang, D. Y., Lin, S. S., & Sun, C. T. (2007). DIANA: A computer-supported heterogeneous grouping system for teachers to conduct successful small learning groups. *Computers in Human Behavior*, 23(4), 1997-2010.

Webb, N. M. (1985). Verbal interaction and learning in peer-directed groups. *Theory into Practice*, 24(1), 32-39.

West, M. A. (2002). Sparkling fountains or stagnant ponds: An integrative model of creativity and innovation implementation in work groups. *Applied Psychology*, 51(3), 355-387.

Worchel, S., Rothgerber, H., & Day, E. A. (2009). Social loafing and group development: when" I" comes last. *Current Research in Social Psychology*, 17 (5).

Zuo, Y., & Serfling, R. (2000). General Notions of Statistical Depth Functions. *Annals of Statistics*, 28(2), 461–482.