How to Assist Tutors to Rebuild Groups Within an ITS by Exploiting Traces. Case of a Closed Forum.

Abstract—Computer Supported Collaborative Learning (CSCL) is a new mode of teaching and one of the popular approaches for learning process. It allows virtual interactions between groups by providing tools such as: chat, internal email and discussion forums. One of the major problems caused by this learning process is the neglect and isolation of learners in groups, and usually is the cause of a heterogeneous group through social, cognitive or emotional ways. The method used is based on the exploitation of traces left on the online learning platform by learners and groups. The data collected from the environment can be observed and exploited in order to build social and cognitive indicators. Our approach is to design a model which assists the tutor to rebuild groups who are not homogeneous in order to prevent their isolation and abandonment. Our model offers the tutor the opportunity to rebuild the groups in an automatic way and based on the characteristics of quantitative indicators of all learners. Our work allowed us to test our algorithm from a functional and technical point of view and also identifies real variables from a collaborative online learning. It also allowed us to evaluate six different indicators proposed for this experiment, showing that they may assist the tutor to rebuild many groups again. The results show us that after the rebuilding groups, there has been a lot of participation in the forum and a considerable number of shares and documents deposited to the forum for each group. This high frequency of interaction between learners, lead them to a fruitful collaboration, and a good quality work at the end.

Keywords—Indicators, Intelligent Tutoring System, Online Collaborative Work, Traces
1 Introduction

In 20 years, distance education has grown from correspondence course on paper to online training in which the information and communications technology (ICT) play a major role. In this case, “Distance Learning” refers to an organization that disrupts traditional habits acquired in lecture halls or classrooms. It is usually done remotely and allows each actor to interact with learning objects, or with other actors, which is a group work that occurs through a learning platform, accessible through the Internet. This online collaborative activity generates a set of traces, which are usually stored in files called logs. In the collaborative context of web 2.0, users are no longer just consumers of content but they are also producers, contributors, and annotators. We make use of the idea that the course of the learners is an information that can be used by tutors to improve resource and guidance for learners [1].

Some works have highlighted the lack of awareness of the activity by the teacher in regards to learner’s activity [2]. This work shows that in an EIAH, the tutor usually gets fewer feedback on the course of the session compared to a conventional learning situation and it is often unable to get a global perception of the virtual class he leads, thus confronting a real identification problem working groups accusing difficulties. Indeed, this element of perception that triggers the decision making of the tutor for further development of the activity (more detailed explanation, complementary exercise, etc.). The distance required from the tutor an extra effort; he must try to understand, decipher, to reconstitute a learner activities puzzle without all the data. An ITS must provides the teacher or trainer accurate information on the individual and collective evolution of learning in order to identify and support individuals facing difficulties avoid their isolation and abandonment. Problem resolution is essentially dependent on the capacity to produce relevant and usable traces of individual or collective learning activity, which interacts with an ITS [3].

The exploitation of interaction traces left by students in the Intelligent Tutoring System and the development of indicators are able to help the tutor to follow up the activities of his learners. Group formation often occurs in a random or according to criteria set by the tutor, so this sometimes leads groups are not homogeneous, hence the neglect and isolation of learners during training. The monitoring of learners in an online learning process constitutes the results of several research projects [4] [5]. Our approach is the use of the traces [6] [7] left by learners to better assist tutors to maintain a balance between working groups. In order to achieve this aim, we will focus particularly on online interactions via a closed educational forum for discussion to each group. This online collaboration would be an appropriate way that will provide more opportunities on the one hand for learners to communicate, share information, interact and virtually produce at the end a quality work. On the other hand, provide teachers to produce and exploit indicators to help and assist them to rebuild the groups are not homogeneous.
2 Related Works

Several research studies have focused on the Intelligent Tutoring System (ITS), especially in remote collaborative work. The literature on ITS has several platforms and tools for the use of individual and collaborative learning, for example: Guéraud & al. [8] and Despres & Coffinet [9] respectively with their "FORMID" and "REFLET" tools, they display the evolution of the work of each student. The tutor can see traces of their student’s activities. Certain workshops have focused on the visualization of the traces from collaborative distance learning platforms. In [10], Mazza & Milani with their "GISMO" (Graphical Interactive Student Monitoring System for Moodle) tool displays the traces in the form of a graph representing the number of entries to the various Moodle resources by different learners. Other studies on ITS have focused on synchronous and asynchronous experiments. In [11], Betbeder & al. with their CoPEAS project (Pedagogical Communication Environment oriented Synchronous Audio) describes a remote collaborative learning experiment mediated by Lyceum [12]. In [13], Greffier & Reffay describes an experimentation exploiting in asynchronous environments such as a forum.

These traces are also used to calculate indicators using quantitative data. In [14], Reffay & Lancieri calculate the cohesion and the centrality in social networks and discussion forums. In [15], Jaillet offers the tutor, via his platform ACOLAD, a tool that provides information on a triplet of the activity (presence, availability, participation). In [16], Santos & al. provide a tool which performs calculations starting from the interactions and the degree of involvement of each learner in training. Other indicators are qualitatively interpreted as Martinez & al. in which they interpret the indicator of the density of the social network by histograms [17]. In [18], Drissi & Amirat propose a new learning style based on online adaptive learning systems (LS-AEHS) to help and guide learners to be well organized during the learning process. As for Ouguengay & al. in [19], they propose a Service Oriented architecture (SOA) based system for tracks retrieving and processing, in order to enhance the effectiveness of the support of users in their interactions within the learning system.

All these studies show the importance of methods and the context of the development of indicators from traces of interactions related to the learning activity.

3 Research problems

3.1 Problematic

In an online collaborative work, working groups are generally formed either a random way or with a compromise between learners or according to criteria set by the tutor, so this sometimes leads groups that are not homogeneous, from which stems the lack of interest of learners towards the learning process. This heterogeneous distribution of learners often leads to the creation of groups with elements having a low level and other groups with elements having a good level. The former groups can act as orchestrators and leaders of the other groups.
To carry out collaborative work, we will study the behavior of learners, the roles they attribute to themselves and the organization they set up to carry out collaborative work. Thus, to progress well in their collective duties, the learners would, therefore, tend to communicate, share information, virtually interact and organize themselves spontaneously to make good progress in their work together. The teacher or trainer must be provided with accurate information on individual and collective learning evolution in order to identify the leaders and support individuals facing difficulties. This will allow them to prevent their isolation and abandonment. Our approach also focuses on the analysis and exploitation of traces left by learners on the platform to assist and help the tutor rebuild groups who are not homogeneous.

To try to understand the dynamics of learning, and efficiently evaluate these learning situations, trace analysis of interaction is exploited by researchers in many systems as evidenced by the state of the art [20] [21]. To this end, we asked the question: How can we draw concrete information from an e-learning process.

This question is the backbone of this article from which several guidelines emerge, namely:

- How data can be extracted from an e-learning platform?
- How traces extracted from the said data can be exploited?
- On what basis can we detect leader learners and learners with difficulties within each group?
- What kind of indicators can be implemented?
- What are the algorithms that can help the tutor to classify learners according to well-chosen criteria?

Thanks to answers to these questions, we will be able to develop a model to advise and assist the tutor to rebuild well working groups following stable and variable characteristics of learners.

3.2 Research Methodology

Research Subjects and Period: We have adopted a methodology based on the study of traces as observation support for the teacher-tutor left in the online learning platform, specifically the discussion forum. In this context, several indicators can be built to observe and operate social and cognitive processes implemented [22]. Indeed, exploiting traces allows building several types of indicators [23]. This experimentation enters into the curriculum of studies in the first year at the ESTS (Ecole Supérieure de Technologie de Salé, Superior School of Technology of Sale) of the Mohammed V University in Rabat. The experimentation was an opportunity to make learners remotely achieve via distance learning platform Moodle [24] a common project, in connection with the Object Oriented Programming in Java. This work lasted six weeks coinciding with the end of the course during the two last week. We chose, this time, to stimulate students to do online work only.

The 85 students involved in this experiment were students of the first year options: Administration of Computer Networks and Software Engineering of the Superior School of Technology of Sale. They were divided into teams of 5 members, 8 groups
of Administration of Computer Networks and 9 groups of Software Engineering. Learners have all followed a traditional course of Object Oriented JAVA programming. The project was set according to the specifications provided by the tutor. The work that these learners have had to carry on the Moodle platform, was recorded in their final note on the “Java OOP” module.

**Tools disposed in platform:** The project was conducted on the Moodle platform. Learners had at their disposal:

- A detailed description of the specifications of the project. (The general context of the project, the functional and technical description, the resources and the time of implementation, etc...).
- Put online resources on the Moodle platform (lectures, articles, Tutorials and corrected practices, Internet links, etc...).
- A structured forum for each group where learners can discuss different tasks and an area of synchronous chat for each group.

**Data collection and analysis:** Traces obtained, usually stored in files called logs and database, they go through various stages before being analyzed and exploited. In our case, data collection was limited to the extraction of information from the Moodle database, log files and the development of quantitative indicators mentioned before, but also to safeguard the jobs posted by learners at the end of their learning processes. In this context, exploiting traces allows us to build several types of indicators [23]:

- **Indicators of the process of the activity:** The indicator “Talk Depth” provides a traceability progress in a forum over time.
- **Indicators relating to the content of the activity:** Indicators related to “Productivity of learners in the forum” and “Productivity groups” are calculated from data collected and shared by each learner and group.
- **Indicators of awareness of the workspace** as the “Degree of the presence of learners” and “Degree of the presence of groups”. These indicators respectively calculate the number of learner connections in a period of time and the number of connections of learners belonging to a group.
- **Indicators relating to the quality of the collaboration** such as “Degree of interactivity of learners in the forum” is an indicator that will calculate the number of messages exchanged within the forum in each group in a given period of time.
- **Indicators of the state of relations between learners:** The indicator “Centrality degree of actors” shows in the form of graphs the email exchanges between members of a group. This indicator helps to identify isolated members, as well as those who dominate the interaction.

### 3.3 Idea and Algorithm

**Idea:** Reconstructing groups remains a problem that has no exact solution or whose solution in a reasonable time remains unknown. The aim is to get adapted solutions to this kind of optimization problem. A metaheuristic is an optimization algorithm able of solving difficult optimization problems (often from the fields of
operations research, engineering or artificial intelligence) for which there is no known more effective conventional method. Metaheuristics behave as search algorithms, trying to learn the characteristics of a problem in order to find out an approximation of the best solution.

There are many different metaheuristics, ranging from a simple local research to global research complex algorithms. However, these methods use a high level of abstraction, which allows them to be adapted to a wide variety of different problems [25].

Let the variable \( P_{i,j,k} \) and \( P \in N \)

- \( i \) corresponds to the indicator \( i = 1, ..., l \) ;
- \( j \) corresponds to the learner \( j = 1, ..., m \) ;
- \( k \) corresponds to the group \( k = 1, ..., n \).

Let \( C_i \) the coefficient of the indicator \( i \). (The coefficient is chosen according to the importance of the indicator).

Let the average of all indicators for a learner defined and calculated by Formula (1).

\[
M_{i,j} = \frac{\sum_{l=1}^{l} (P_{i,j,k} \times C_i)}{\sum C_i}
\]  

Let the general average of the indicators for each group defined and calculated by Formula (2).

\[
MG_{k} = \frac{\sum_{j=1}^{m} M_{j,k}}{m}
\]  

Let the arithmetic average and standard deviation of the groups defined and calculated by Formula (3) and Formula (4)

\[
\mu = \frac{\sum_{k=1}^{k} MG_{k}}{n}
\]

\[
\sigma = \sqrt{\frac{\sum_{k=1}^{k} (MG_{k} - \mu)^2}{n}}
\]

The best algorithm that can solve this problem in order to find a balance in groups is the descent algorithm with a neighborhood 2-opts. The idea of this algorithm is to swap learners between groups until the standard deviation of the average of the groups tends to zero. The number of neighborhood in this method is \( n(n+1)/2 \) and \( O(n) = n^2 \).

We consider the distribution of learners as follows:

Solution \( S = (a_1, ..., a_n) \) such as \( (a_1, ..., a_n) \in MG_k \) and \( n \) is the number of learners.

We consider the neighborhoods \( V(S) \) of a solution \( S \) as follows:

\[
V(S) = \{a_1, ..., a_{i-1}, ..., a_i, a_{j-1}, ..., a_j, a_{j+1}, ..., a_n\}
\]
For each permutation, we calculate the average for each group $M_{G_k}$ and the standard deviation $\sigma$.

Once all the permutations are finished, we look for the solution which has the smallest deviation and apply the same algorithm to this one. If we do not find a solution that improves the standard deviation found out, we suppose that this is the appropriate constitution of groups. Otherwise, we repeated the same treatment until there is no longer a standard deviation, which improves the previous solution.

**Algorithm:** The algorithm we chose that can solve this problem in order to find a balance in groups is the descent algorithm with a neighborhood 2-opts (Local search).

**The Algorithm**

- Generate an initial standard deviation $\sigma$.
- iteration loop.
  - Generate the neighborhood of $V(\sigma)$ of $\sigma$.
  - Search $\sigma'$ that $\in V(\sigma)$ such as $f(\sigma') = \min f(\sigma), \sigma \in V(\sigma)$
  - If $f(\sigma') < f(\sigma)$ then $\sigma := \sigma'$ and go to iteration 1

**3.4 Research Results**

**Evolution of groups:** The first crucial remarks that are found for all groups is that during the first week, learners often log in to the platform and participate in the discussion forum following the two indicators “Degree of the presence of learners” and “Degree of interactivity of learners in the forum”. This period had not much sharing and document submissions on the forums of each group as we have seen from the Productivity of learners in the forum indicator. The second week we have seen a large number of message exchanges on the forums of each group as shown in Table 1 and described in Figure 1. Thus we could detect people who can play the role of leaders in the group based on their presence in the platform and their interests in the discussion forum.

Based on the indicators: “Productivity of learners in the forum” and “Degree of interactivity of learners in the forum” it was found that learners as: (Mouzakki & Kachtouli, Group 1); (Zine & Jbili, Group 2); (Bouaziz, Group 5); (Iguichi & Setti, Group 7); (Laariche, Group 9); dominate and monopolize the interactions (with a percentage of 70%) in their groups and share a lot of information with their colleagues. These students they play well the role of an orchestrator in their groups.

Figure 2 shows us the performance of all groups by calculating the average of quantitative indicators cited above. The histogram indicates that there is heterogeneity within some groups and especially for groups (3, 4, 5 and 8).
Fig. 1. Indicators relating to « The productivity of learners » and « The degree of interactivity of learners » for the first three weeks.

Table 1. Number of messages exchanged and documents shared for each group during the three first weeks

<table>
<thead>
<tr>
<th>Groups</th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>2</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Group 2</td>
<td>1</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Group 3</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Group 4</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Group 5</td>
<td>7</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Group 6</td>
<td>7</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>Group 7</td>
<td>10</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Group 8</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Group 9</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Fig. 2. Average of quantitative indicators in Software Engineering groups in the forum platform

176 http://www.i-jet.org
The re constitution of groups: By the end of the second week, we made a new reconstruction of the groups considering the heterogeneity of some groups. The constitution was done in an automatic manner according to our algorithm [26]. Each group will have an average following quantitative indicators that we have exploited. The idea of this algorithm is to swap the learners between groups until the standard deviation of the average of the group tends to zero. Once the new groups are built, we observed the evolution of and collaboration between learners.

The first observation that we made during the last two weeks is there has been a lot of participations in the forum and a considerable number of shares and documents deposited to the forum for each group as shown in Table 2 and described in Figure 3. We also noticed that those leaders have played well their orchestrator roles within their groups. This observation is the result of a comparison of a few learners before and after reconstitution of groups.

Debate on the Research Results: The study we presented shows that learners were forced to work remotely via an online learning platform. In the absence of a tutor who can direct and guide them, they have the challenge and were able to manage this new learning situation. These learners have used virtual interactions and exchanges in a discussion forum to ask for help, explain, share and learn in order to provide a final work.

This asynchronous collaborative learning mode, through a discussion forum, can save the texts of the various participants in a discussion and to preserve them permanently. Allowing the return to these messages at any time for learners who have reactions, clarifications, answers and summaries, can mark their process of understanding and develop their own thoughts [27].

Some learners may produce nothing if there is an absence of a cohesion in the group. These include the problems of the affective dimension between learners, convergence problems and those of perception of the group. All these factors can nega-
tively affect the process of discussion, sharing and collaboration. (Group Case 8).

Figure 4 shows us the performance of all the new groups by recalculating the average of quantitative indicators cited above. The histogram indicates that there is an improvement and an evolution compared to the first constitution of groups, it shows also that there is some kind of stability and homogeneity between groups.

Our research on online collaboration and exploitation of traces of learners/groups in a discussion forum leads to grow our studies further and ask more questions. How can one know the reliability and quality of discussions exchanges between learners? How to make the discussion forum that will allow a friendly interactions leading to sharing and collaboration? Can we say that learners prefer exchanges by email or via Skype (Case in groups 1 and 3) than through the forum? [28][29]. Several questions we will try to answer in future empirical research, including structuring the forum by categorizing it.

Table 2. Number of messages exchanged and documents shared for each group during the three last weeks

<table>
<thead>
<tr>
<th>Groups</th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>4</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Group 2</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Group 3</td>
<td>2</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Group 4</td>
<td>4</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Group 5</td>
<td>5</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Group 6</td>
<td>6</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Group 7</td>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Group 8</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Group 9</td>
<td>3</td>
<td>8</td>
<td>11</td>
</tr>
</tbody>
</table>

4 Conclusion

Our work aims to find a cohesion and a significant motivation in a group through an online learning platform. This homogeneity can only be built if learners have the prerequisites for remote collaborative work. Forced to work in groups and in a discussion forum, learners behave the same way as in-face learning situations and have the same behavioral patterns as those adopted in the conventional group (Planning issues, answers, clarification, and comments, etc...). The presence of learners, the frequency of interactions, the number of documents shared within a forum can lead to fruitful collaboration, which is a synonym of good quality work. Such indicators can detect learners and groups who face difficulties and give an idea on learners who can play the role of a leader in a group.

This work allowed us to test our algorithm from a functional and technical point of view and also identifies real variables from a collaborative online learning. It also allowed us to evaluate six different indicators proposed for this experiment, showing that they may assist the tutor to rebuild many groups again. One of the perspectives of this work concerns the integration of other more advanced indicators and especially
indicators of an effective dimension of learners. This dimension consists of developing a sociogram through a questionnaire for learners to get an idea of the emotional state, feelings and learners affect. Other indicators can be modified by categorizing the messaging forums and email in order to provide better visibility and tutor assistance to help rebuild the groups that face difficulties.

Fig. 4. Indicators relating to “The productivity of learners” and “The degree of interactivity of learners” for the last three weeks.

5 References

Paper—How to Assist Tutors to Rebuild Groups Within an ITS by Exploiting Traces. Case of a Closed…


6 Authors

Salihoun Mohammed was born in Rabat, in 1986. He is a PhD student at Mohamed Mohammed V University in Rabat, Morocco. He received a Master’s degree in Computer Science from Mohamed V University in Rabat in 2011. His research interests are E-learning, Tracking and Software Engineering.

Guerouate Fatima is a Research professor at LASTIMI laboratory and Professor in the Department of Computer Sciences at the Superior School of Technologies of Salé, Mohammed V University in Rabat, Morocco. She is interested in E-learning and Serious games.

Naoual Berbiche is a Research professor at LASTIMI laboratory and Professor in the Department of Computer Sciences at the Superior School of Technologies of Salé, Mohammed V University in Rabat, Morocco. She is interested in Models transformation, systems interoperability, computer network security and web application security.

Sbihi Mohamed. He received the Master degree and the Doctorate in Automatic and Information Processing at Mohammed V University in Rabat. He is interested in Image processing and Data analysis. He is the author or coauthor of some papers in international conferences and journals.

Article submitted 10 December 2016. Published as resubmitted by the authors 27 January 2017.