

UPCommons

Portal del coneixement obert de la UPC

<http://upcommons.upc.edu/e-prints>

Aquesta és una còpia de la versió *author's final draft* d'un article publicat a la revista *Expert systems with Applications*.

URL d'aquest document a UPCommons E-prints:

<http://hdl.handle.net/2117/102091>

Paper publicar / *Published paper*:

Gaudioso, Elena, Montero, Miguel, Talavera, Luis and Hernandez-del-Olmo, Felix (2009) Supporting teachers in collaborative student modeling: a framework and an implementation. *Expert Systems with Applications*, Volume 36, Issue 2, Part 1, March 2009, Pages 2260-2265, ISSN 0957-4174, DOI:10.1016/j.eswa.2007.12.035.

Elsevier Editorial System(tm) for Expert Systems With Applications

Manuscript Draft

Manuscript Number: ESWA-D-07-00565

Title: Supporting teachers in collaborative student modeling: a framework and an implementation

Article Type: Full Length Article

Keywords: Student model; collaborative user modeling; intelligent learning environments; educational data mining

Corresponding Author: Elena Gaudio, PhD

Corresponding Author's Institution: UNED

First Author: Elena Gaudio

Order of Authors: Elena Gaudio; Miguel Montero; Luis Talavera; Felix Hernandez-del-Olmo

Supporting teachers in collaborative student modeling: a framework and an implementation

Elena Gaudioso^a Miguel Montero^b Luis Talavera^c
Felix Hernandez-del-Olmo^a

^a*Artificial Intelligence Department, Universidad Nacional de Educacion a Distancia. Juan del Rosal, 16, 28040 Madrid Spain, {elena, felixh}@dia.uned.es*

^b*ICT Consultant in the Teacher Resource Center. Avenida de España 12, 02002 Albacete Spain, mmontero@jccm.es*

^c*Dept. Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Jordi Girona 1-3, 08034 Barcelona, Spain, talavera@lsi.upc.edu*

Abstract

Collaborative student modeling in adaptive learning environments allows the learners to inspect and modify their own student models. It is often considered as a collaboration between students and the system to promote learners' reflection and to collaboratively assess the course. When adaptive learning environments are used in the classroom, teachers act as a guide through the learning process. Thus, they need to monitor students' interactions in order to understand and evaluate their activities. Although, the knowledge gained through this monitorization can be extremely useful to student modeling, collaboration between teachers and the system to achieve this goal has not been considered in the literature. In this paper we present a framework to support teachers in this task. In order to prove the usefulness of this framework we have implemented and evaluated it in an adaptive web-based educational system called PDinamet.

Key words: Student model, collaborative user modeling, intelligent learning environments, educational data mining

Supporting teachers in collaborative student modeling: a framework and an implementation

1 Introduction

Traditionally, the main goal of adaptation in educational systems is to guide the students through the course material in order to improve the effectiveness of the learning process. One of the most prominent example of this fact is the research area of Intelligent Tutoring Systems (ITS) [10]. The main goal of a ITS is to adapt the teaching according to the students' individual skills, knowledge, and needs, and give personal feedback just in time [3].

This adaptation is done on the basis of the information stored in the so-called student models. The underlying mechanisms to construct and maintain the student models are not so different from those used for user modeling [6] in user-adaptive systems. One requirement widely admitted is that user models should be scrutable [8]. This requirement is two fold. On the one hand users should have the opportunity to inspect their user models to understand the system behaviour. On the other hand, the user-adaptive system would need the cooperation of the user to provide some information that can not be obtained otherwise. When the user cooperate with the system to construct and mantain the user model it is known as *collaborative user modeling*.

In educational systems collaborative student modeling has an added value. Collaborative student modeling is used to make learners inspect and modify their own student models. A more accurate model may thereby be obtained, and learners may reflect on their beliefs while constructing their model [2]. This approach is also used to collaboratively assess the course, by the use of open models of the student's knowledge. The effectiveness of open student modeling on learning has been already reported in the literature [12]. However, when systems are used in a classroom, students are not just supposed to learn on their own, but a teacher acts as a guide and facilitator through the learning process. Teachers need to understand and evaluate the activities of the students while using the educational application and this can only be accomplished by analyzing student interactions and performance. It appears obvious that the knowledge gained through this monitorization can be extremely useful to interactively modify student models or, in other words, that teachers should have a key role in a collaborative user modeling process.

To our knowledge, the collaboration between teachers and the system to

achieve student modeling has not been considered. Nevertheless, it is consistent with the goal of advocating the role of the teacher as an essential actor when applying adaptive educational systems in classrooms.

In this paper we present a framework to support teachers in collaborative student modeling. To illustrate our proposal we have implemented it in a web-based adaptive educational system for physics teaching in secondary education called PDinamet.

The remainder of the paper is organized as follows. In Section 2 we present the proposed framework. Its implementation in Pdinamet is described in Section 3. Next, in Section 4 a description of the experiments we have carried out to validate our approach is presented. To conclude we present in Section 5 our conclusions and lines of future work.

2 A framework for supporting teachers in collaborative student modeling

There are situations in the development of user-adaptive systems in which the only way for the system to gather the required information about the users is to engage them in the process of user modeling and to collaborate with the user in gathering the information required. This approach is known as *collaborative* [1] or *cooperative* [7] user modeling.

A user-adaptive system can be viewed as an user-modeling module working along with a particular application through three stages [5,1]. First, it collects data about the user interaction with an application. Using this data, the system builds a user model by performing some type of learning and/or inference on the basis of this information. Finally, in the process of user model application an adaptation of the application behavior is determined in order to better fulfill the user goals.

According to these stages, [1] differentiates three approaches to involve the user in the user modeling process. First, users can directly provide the data required for the user modeling mechanism. Second, the information in the user model can be updated directly by the information received from the users. Finally, users can make the desired adaptation themselves, directly showing the system what they would like in a given context.

In educational systems, collaborative student modeling is used to make learners inspect and modify their own student models and to collaboratively assess the course. There are several possibilities for course assessment. In [14], for example, an off-line evaluation scenario for collaborative learning tools

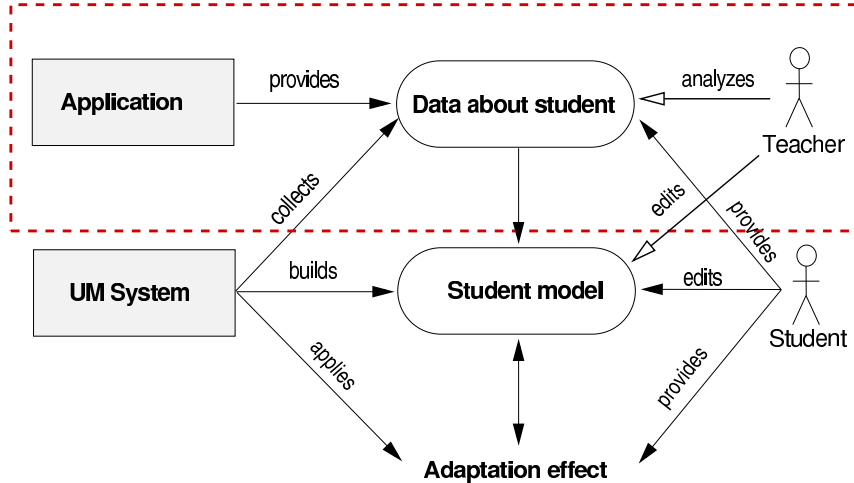


Fig. 1. Collaborative student modeling in educational systems.

is suggested. In this case, assesment can be performed by the system or a human evaluator that can intervene in the process to alter students behavior.

Figure 1, adapted from [1], shows the general collaborative student modeling process formulated for educational systems. An educational application provides a series of elements for students and provides data to the user modeling system (named UM System in the Figure), which collects the data and builds and applies the student model to provide some kind of adaptation on the original application (not shown in the figure). Note that both, students and teachers, can influence the process. The three possibilites for involving users mentioned above are illustrated but, in the case of teachers, they only intervene by analyzing the application data and modifying student models making use of the knowledge obtained.

The goal of collaborative user modeling systems is to develop strategies to improve the communication process so that the machine and the user can work effectively together [7]. In our case, this translates into providing teachers with analytical tools that can assist in obtaning knowledge from the large amount of data generated by the interaction of students with the application. We have outlined elsewhere [15] that the general management cycle suggested in bears many resemblances with the typical data mining cycle, namely, data collection and preprocessing, building an analytical model, evaluation and interpretation and deployment, incorporating the model into another system for further action. If we look at the top of figure 1, that illustates which parts of the process are influenced by the inclusion of teachers in the collaborative student modeling process, it is easy to see that it also fits perfectly with the formulation of a data mining task.

The realization of these ideas is illustrated in Figure 2 where a general framework to support teachers in collaborative student modeling through data min-

ing is depicted. Figure 2 b) shows the typical setup for an adaptive educational system, including collaborative modeling by involving students. The left part presents the proposed addition in the form of a typical data mining cycle. First, the teacher defines an analysis task to perform (1), which usually would turn into choosing the appropriate subset of data for the analysis and a particular algorithm that fits the goals of the analysis. Note that in other user-adaptive systems the required data and inferences may be known beforehand. This is not usually the case in educational environments because teachers may have different goals in mind depending on the context, so that data and algorithm selection may vary from case to case. When the data mining model is built (2), some tool for interpret and evaluate the resulting patterns is provided. Finally, in the deployment step (3), teachers can modify the models of adaptation in the system, typically student and pedagogical models.

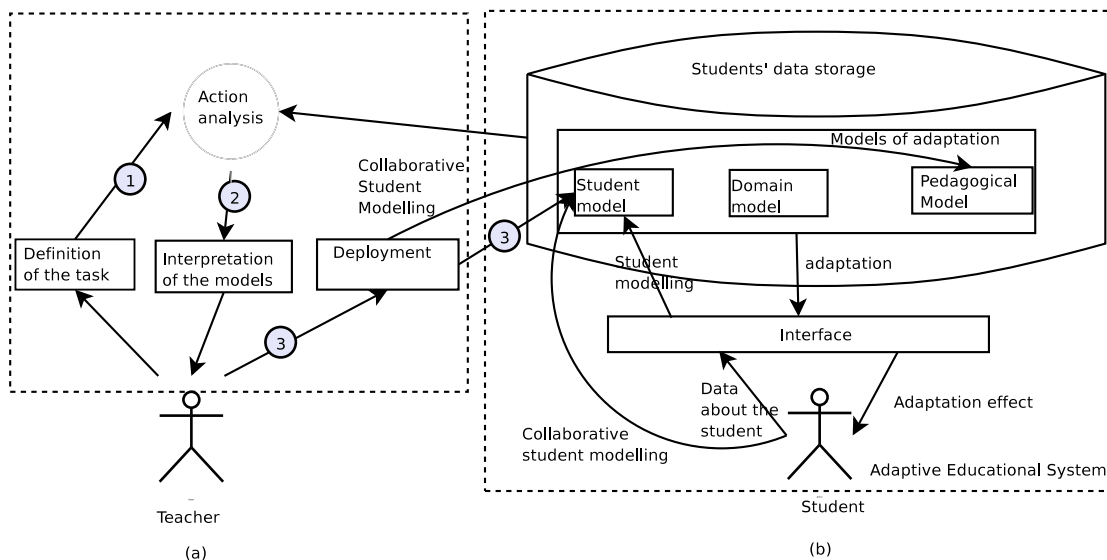


Fig. 2. General Framework to support teachers in collaborative student modeling.

It is important to note that specific implementations of this framework should not only pay attention to the analytical technology -data mining algorithms- used, but also have to provide tools for the full cycle, including the interpretation and deployment stages. Data mining results shown in raw form can be difficult to interpret so that visualization tools may provide better insight to users that are not familiar with analytical concepts [9]. The deployment of the knowledge obtained with the analytical models may not be a trivial task and an effort should be made to try to integrate this step into the particular application seamlessly.

3 Collaborative student modeling in PDinamet

PDinamet is a Web-based adaptive Learning system directed to the teaching of the Physics in secondary education [13]. PDinamet is based on the following elements:

- **Domain model.** In PDinamet there are different types of learning resources, such as workshops, exercises generators or theory pages. Each resource is represented by a set of characteristics describing aspects such as learning goals or level of difficulty. Teachers are allowed to add new resources. Each resource is associated with an indicator so that a set of indicators corresponds with a concept. The relationships between different learning resources and concepts make the domain model of the system. When a learner has achieved a set of indicators corresponding with a concept PDinamet assumes that this concept has been assimilated.
- **Student model.** In PDinamet we have considered the student model as an overlay model [4]. Thus, for each concept in the course and for each student we store if the user have learned or not the concept. The student model in PDinamet also contains personal data, academic data, computer skill level and background knowledge. The latter comprises several attributes with information about both the knowledge level that the students demonstrate in a previous test and what they consider they already know. The student model is enriched with sets of recommended resources that are selected according to the specific knowledge level and activity of each student.
- **Pedagogical model.** It includes the information needed to guide students through the course and it is implemented as a recommender that supports selection of the appropriate learning resources depending of each particular context. This is accomplished by means of a set of predefined rules stored in the knowledge base of a Prolog program. The rules are of the form `recommended-item(studentID, resourceID)` and test wether a particular student satisfy a set of given pedagogical requirements defined in terms of the student current knowledge and characteristics of the available resources.

The models described above are encoded in a relational database except the pedagogical model and the recommended items of the student model which are respectively encoded in the knowledge base and as facts of the Prolog program that acts as a recommender.

In order to be able to modify student models in a collaborative user modeling environment, teachers need to assess the course by analyzing the interactions with the application. Following the proposed framework, support for this task in PDinamet is provided by means of a data mining tool that characterizes patterns of student behavior. To this end we focus on unsupervised methods which are descriptive in nature. Teachers may confirm or deny intuitions about

the course and, possibly, particular cases of interest.

The first step in building a data mining model is to choose the data that is going to be used for the analysis. This is particularly important in this case, since teachers may select different subsets of data for assessing different aspects of the course. The input data considered is described in table 1. They include both background data and interaction data, including access data, use of resources and grades. Numerical data are preprocessed with a discretization procedure in order to simplify the interpretation of the results.

Background data	Personal data , academic data, computer skill level, background knowledge, learning style.
Access data	Total number of visits to the system, Number of sessions, Number of short sessions, Number of long sessions, total time the student has really worked with the system.
Use of learning resources	Percentage of use of the different learning resources , percentage of learning resources that have been correctly solved.
Performance	Grades obtained by the students in different concepts of the course

Table 1
Input data available for teachers in PDinamet for building data mining models.

Using the selected data, the analysis tool will construct the analytic model. In the current implementation, we employ clustering techniques to discover groups of learners with common behavioral trends to help teachers inspect which learners have problems and should receive some reinforcement. There exists a large number of clustering algorithms and the choice depends on the particular application. For our purposes, we choose model-based clustering which is an approach that has gained wide popularity in the clustering literature [11] and it is provided by the Weka machine learning tool[16].

When building data mining models, a teacher is first presented with a data selection screen showing the available attributes grouped by types as shown in Table 1 to simplify the process. A particular attribute can be selected as an external characterization feature in order to help teachers to interpret the models. This adds a different view to the clusters that makes possible to assess if the discovered behavioral patterns are correlated with known goals (for example, if a student has failed the course). This attribute is external in the sense that it is not used during the clustering process. Next, the number of clusters can be selected or let the system to automatically approximate the best number given the data. Once a model is built, a screen showing a summary of results is shown.

To support the inspection of the results we have implemented a simple but useful visualization tool that shows the distribution of the values for an attribute among the clusters. Initially, the distribution of the external feature or class is shown to provide an initial interpretation of results. As depicted in Figure 3, for each cluster and for each attribute value, bars with different color segments are shown representing the number of students exhibiting a particular value for the attribute. From this initial view, teachers can proceed to inspect the distribution of the rest of attributes by selecting them from a list.

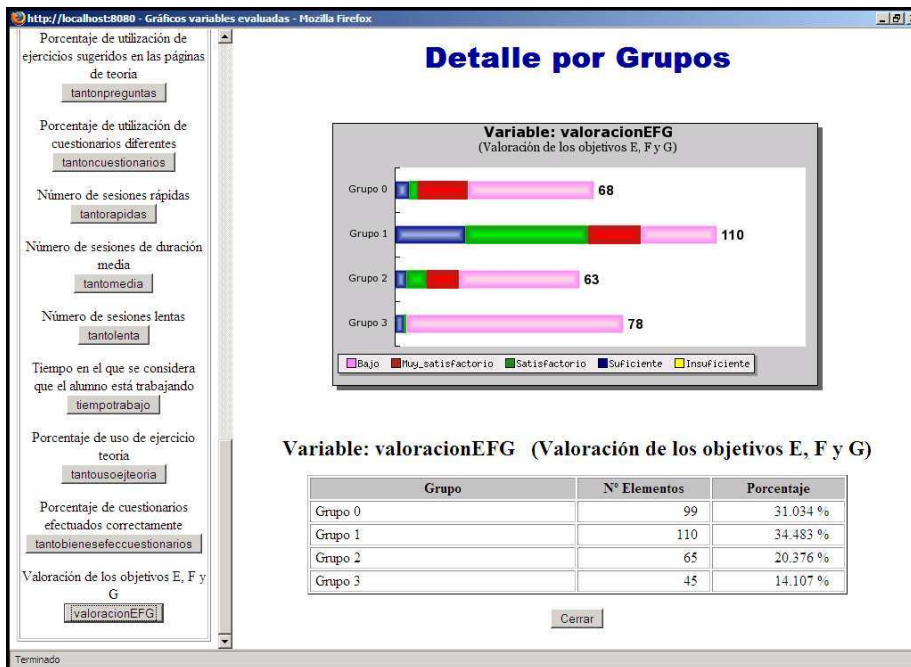


Fig. 3. Diagram bars with the visualization of the results (in Spanish)

The features presented so far cover the data preparation and model building and evaluation steps of the data mining process, but in order to close the loop in the context of the framework proposed, teachers also have to be able to modify student models according to the insight gained from the analysis. When a teacher clicks on one of the portion of the bars of the graph the list of students included are shown. Then the teacher has the chance of assigning several resources to be recommended only to those students. Once one or more resources are selected, each student model is updated adding the Prolog predicates that allow to the recommender make the recommendations of these resources.

4 Experiments and Evaluation

To test the tool implemented, PDinamet has been tested in an experiment in several secondary schools with about 300 students from 15 different classes that used the learning environment for a period of 2 academic years. The subject area was Dynamics.

In our experiments we have considered the three first sets of features in Table 1 leaving the data about the performance as external characterization features. In Figure 3 we show a visualization of the clusters obtained and their correlation with the final grade. The grade has been discretized in 4 bins, namely, *very high*, *high*, *medium* and *low*. It can be seen in Figure 3 that the last cluster there are a great number of students with a final grade of *low*. At first sight, these students need help from the teacher since their performance is not good. To get better insight into what characterizes these students, the teacher can further explore the clusters by inspecting the distribution of the rest of attributes in the left pane of the Figure 3. In this particular case, these students are too much confidence in their initial knowledge in the subject and thus they have visited few theory pages and most of their sessions are very short (perhaps only for gaming). Teachers may want to motivate them to study harder by recommending new types of learning resources besides theory pages.

In Table 2, we show a summary of the profiles we have obtained for each cluster described both in terms of the input features and the value of the final grade. Note that unlike before, for this summary we have considered only if the student has passed or not. Thus if a student has a final grade of low it is considered that she has failed and she has passed otherwise. They can be used by tutors, for example, to identify some effects derived from group behavior.

We have evaluated this tool with the teachers of the students that have participated in the experiments of PDinamet (eight teachers). Besides the module presented in this section, the complete tool presented to teachers includes a reporting tool that offers information about how individual students and group of students have performed in the course (in terms of rates of study or grades). At the end of the course they filled a questionnaire to help us evaluate the module. All teachers in the experiment have used this tool for the assessment of their course and that proves its usefulness. The items evaluated were the satisfaction of the teachers with the system, the level of use, the usability and the satisfaction with the clusters presented. All these items have a score above 16 (the maximum score is 25). The strongest point was the satisfaction of the teachers with the clusters presented. The weakest point in the results was the usability that is probably because the necessary expertise to understand the whole process (we expect to improve this aspect in the future).

Cluster (%)	Discriminant	External
1 (0.30)	High confidence in their background knowledge, they really have good background knowledge, good knowledge in basic concepts, low number of visits to the system, the sessions have an average duration and the time the students are working in the system is high	Mixed (0.37/0.63)
2 (0.26)	They are considered themselves diligents, not confident with their own background knowledge but they really have good background knowledge, high number of visits mainly to theory pages, they have long sessions working in the system	Pass (0.76/0.24)
3 (0.20)	High number of misconceptions in their background knowledge, high number of visits to the system mainly to theory pages and the exercises proposed	Mixed (0.35/0.65)
4 (0.24)	High confidence in their background knowledge, low number of visits, high number of short sessions, low number of visits to theory pages	Fail (0.95/0.05)

Table 2

Clusters including discriminant features, and an external profiling feature (pass/fail).

5 Conclusions and Future Work

In this paper we have presented a framework to support teachers in collaborative student modeling and an implementation example into a particular educational application, PDinamet. The results obtained in PDinamet have shown the usefulness of clustering techniques to obtain and characterize groups of students with different profiles. Exploration of results allow students that might require special attention to be identified and characterized in terms of their interaction with the system and other information available.

Our framework aims to make the data mining results useful in practical situations, so that analytical technologies have to be completed with some strategy to interpret and evaluate the results and an appropriate interface. Moreover, once the results are validated they must be actionable and have some impact in the user modeling system. Our experience shows that simple visualization

techniques and interactive interfaces allow users to explore the clustering results and obtain good insight. We have closed the loop by linking these results to the user modeling system by updating student models that can be used by the recommender in further interactions. That said, although the preliminary evaluation with teachers have proven positive, it also reflects the need to develop both, postprocessing techniques to facilitate the interpretation of analytical models and the usability of the interfaces to support this process.

In PDinamet, every interaction of the learner with the recommender is traced, so, in future work, we plan to check if the recommended learning resources have been useful for the learner by checking if the learner improves their learning results. When this happens, a special event could be sent to the recommender to update its knowledge base by removing the Prolog predicates associated with that resource.

References

- [1] P. Brusilovsky. Methods and techniques of adaptive hypermedia. *User Modeling and User Adapted Interaction-Special issue on adaptive hypertext and hypermedia*, 6(2-3):87-129, 2003.
- [2] S. Bull. 'do it yourself' student models for collaborative student modelling and peer interaction. In B.P. Goettl, H.M. Halff, C.L. Redfield, and V.J. Shute, editors, *Proceedings of International Conference on Intelligent Tutoring Systems*, pages 376-185, 1998.
- [3] W. Hamalainen, T.H. Laine, and E. Sutinen. Data mining in personalizing distance education courses. In C. Romero and S. Ventura, editors, *Data Mining in e-learning*, pages 157-171, 2006.
- [4] P. Holt, S. Dubs, M. Jones, and J.E. Greer. The state of student modelling. In J.E. Greer and G.I. McCalla, editors, *Student modelling: They key to individualized knowledge-bases instruction*, pages 3-38, 1994.
- [5] A. Jameson. Adaptive interfaces and agents. In J.A. Jacko and A. Sears, editors, *Human-Computer interaction Handbook*, pages 305-330, 2003.
- [6] R. Kass. Student modeling in intelligent tutoring systems: Implications for user modeling. In A. Kobsa and W. Wahlster, editors, *User Models in Dialog Systems*, pages 386-410. Springer, Berlin, Heidelberg, 1989.
- [7] J. Kay. The um toolkit for cooperative user models. *User Modeling and User Adapted Interaction*, 4(3):149-196, 1995.
- [8] J. Kay. An exploration of scrutability of the user models in personalised system. *Engineering Personalised Systems*, 2002.

- [9] R. Mazza and C. Milani. Exploring usage analysis in learning systems: Gaining insights from visualisations. In *Proceedings of the 12th International Conference on Artificial Intelligence in Education (AIED 2005)*, Amsterdam, The Netherlands, 2005.
- [10] D. McArthur, M. Lewis, and M. Bishay. The roles of artificial intelligence in education: Current progress and future prospects. *RAND DRU-472-NSF.*, 1993.
- [11] M. Meila and D. Heckerman. An experimental comparison of model-based clustering methods. *Machine Learning*, 42(1/2):9–29, 2001.
- [12] T. Mitrovic. Evaluating the effects of open student models on learning. In P. De Bra, P. Brusilovsky, and R. Conejo, editors, *Proceedings of Adaptive Hypermedia and Adaptive Web-Based Systems*, pages 296–305, 2002.
- [13] M. Montero and E. Gaudioso. P-dinamet: A web-based adaptive learning system to assist learners and teachers. In C. Ghaoui, M. Jain, V. Bannore, and L.C. Jain, editors, *Knowledge-Based Virtual Education*, pages 23–48, 2005.
- [14] A. Soller, A. Martinez, P. Jermann, and M. Muhlenbrock. From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *International Journal on Artificial Intelligence in Education*, 15(4):261–290, 2005.
- [15] L. Talavera and E. Gaudioso. Mining student data to characterize similar behavior groups in unstructured collaboration spaces. In *Proceedings of the Workshop on Artificial Intelligence in CSCL. 16th European Conference on Artificial Intelligence, (ECAI 2004)*, pages 17–23, Valencia, Spain, 2004.
- [16] I.H. Witten and E. Frank. *Data mining: Practical machine learning tools and techniques*. San Francisco, 2005. Morgan Kaufmann.