Redes Bayesianas para detección de roles de equipos en aprendizaje colaborativo soportado por computadoras

Recibido: 15 Febrero 2015 – **Revisado:** 30 Abril 2015
Aceptado: 30 Mayo 2015 – **Publicado:** 30 Julio 2015

José Balmaceda
Facultad de Ciencias Exactas, Universidad Nacional del Centro de la Pcia. de Bs. As., Campus Universitario, Paraje Arroyo Seco, Tandil, Argentina.
jmbalmaceda@gmail.com

Silvia Schiaffino
Instituto de Investigación SISTAN (CONICET – UNCPBA), Campus Universitario, Paraje Arroyo Seco, Tandil, Argentina| silvia.schiaffino@isistan.unicen.edu.ar

Daniela Godoy
Instituto de Investigación SISTAN (CONICET – UNCPBA), Campus Universitario, Paraje Arroyo Seco, Tandil, Argentina.
daniela.godoy@isistan.unicen.edu.ar

Resumen: El trabajo colaborativo soportado por computadoras permite a los estudiantes que se encuentran en lugares remotos trabajar de manera conjunta en el mismo entorno virtual y permite la comunicación de ideas e información entre los integrantes del grupo. Sin embargo, como no todos los estudiantes son iguales, es importante estudiar las características de éstos para construir grupos de trabajo más productivos. La teoría de roles de equipo posibilita obtener buen desempeño en los equipos de trabajo considerando habilidades individuales, combinando las falencias de cada rol con las fortalezas de los otros. Generalmente, las personas tienen que completar extensos cuestionarios para poder determinar sus roles de equipo. En este trabajo, se propone un método alternativo para realizar esta detección a través de un sistema de aprendizaje colaborativo y a partir de la utilización de la técnica de Redes Bayesianas.
Palabras clave: roles de equipo; trabajo colaborativo soportado por computadoras; redes Bayesianas.

Abstract: Computer-supported collaborative learning allows students who are in different places to work together in the same virtual space, and supports the communication of ideas and information among learners. However, as not all students are identical, it is important to study users' characteristics to build more productive teams. Team Roles Theory allows obtaining very good team performance taking into account individual skills, combining the weaknesses of each role with the strengths of others. Originally, people have to complete extensive questionnaires to determine their team role. In this work we propose an alternative method to make this detection through a collaborative learning system and by using a Bayesian Network.
Key words: team roles; computer supported collaborative learning; Bayesian networks.

1. INTRODUCTION

In the last years, the increase of accessibility to the Internet has augmented the use of e-learning in different institutions. One of the most widely used types of applications in e-learning are computer supported collaborative tools. Computer-supported collaborative learning (CSCL) allows students who are far apart to collaborate on-line, and supports the communication of ideas and information among learners who collaboratively access to information and documents. In this context, students have to carry out several types of activities. We will focus our attention on group activities, such as solving a given exercise proposed by the teacher in groups or teams.

Achieving an effective collaboration among the participants of a team work is a difficult problem. The productivity of groups or teams can be determined by the way in which group members work collaboratively. In this context, the collaborative skills and particular features of team members will be a key factor in the success or failure of the underlying collaborative learning task. Not all students are alike; they have different behaviors, different knowledge and different abilities to carry out a given task. These differences might be a source of conflict and if they are not managed properly they might diminish team performance and affect learning in a negative way. Team Role Theories have emerged as an alternative to cope with this problem. These theories study the way in which people with different characteristics should be merged or combined into groups in order to form productive and efficient teams. Most of the proposed models require filling extensive questionnaires, which can be a tedious activity that can take a long time, and people could make mistakes.

This article presents an automatic intelligent technique for team role detection in CSCL. This technique does not require that students complete extensive questionnaires to determine their roles and is carried out without users noticing it. CSCL applications register huge amounts of data about student-application interaction and student-student interaction. These data constitute a valuable source of information about how team members work. This information enables us to build, using a Bayesian network (BN), a student model that describes each team member behavior and preferences, and can be used to determine the roles that best fit a given student. Knowing each student's team role could be a very useful tool to make more successful learning groups, taking into account individual skills, combining the weaknesses of each role with the strengths of others.

The rest of the article is organized as follows. In Section 2 we present some background information about team roles and related work on team roles detection in collaborative systems. In Section 3 we describe the technique proposed in this work. In Section 4 we present the results obtained when evaluation the precision of our proposal. Finally, in Section 5 we present our conclusions and future work.

2. BACKGROUND: TEAM ROLES

The concept of role was defined as a group of behavioral patterns expected and attributed to someone that occupies a certain position in a social unit (Aguilar et al. 2007, Belbin 1993). Several models and theories have emerged that study how the different roles contribute to the group work, and propose the different roles that people can take in a work group. Belbin (1981) was the first researcher who proposed a team role theory, which is one of the most widely known in the area of group formation. Belbin explains that a team role is identified as a behavioral pattern that is characteristic of the way in which a team member interacts with others to facilitate group progress.

In the literature we can find several team roles models, apart from Belbin, such as Mumma (1992), MTR-I (http://www.profiles-r-us.com/mtri), Insights (http://www.insights.com/), among others. After a deep bibliographic review, we decided to use in our work the model proposed by Mumma (1992). This model coincides with Belbin's in several aspects. Mumma's model is simple to apply since it provides a questionnaire that enables everyone to determine his/her favorite roles. We describe the main features of this model in the following section.

2.1. Mumma's Team Role Model

Frederich Mumma developed a team role theory composed by 8 roles. He also developed the Personal and Team Roles Profile based on the Team Work cycle (Mumma 1994). This theory is based on studies of Bales and Strodtbeck (1951) that describe the different phases that a team suffers to solve a problem. Mumma defines four phases in the team work life cycle. Each phase consists of two defined team roles.

- Phase 1: Initiation. It occurs when a task is defined. Leader and Moderator team roles are important in this phase. The leader inspires and motivates the team members. The moderator matches the resources to the task at hand.

- Phase 2: Identification. It allows the team to identify alternative methods to perform a task such that needs could be fulfilled. Creator and Innovator team roles are important in this phase. The creator identifies original ideas to approach a task along with alternatives. The innovator identifies opportunities to use the various resources in the firm.

- Phase 3: Elaboration. This phase covers the elaboration of ideas invented from the ideation phase. The objective of this phase is to make the ideas work properly. Improper elaboration can cause conflicts with people, schedules, budgets and other resources. Manager and Organizer team roles are important in this phase. The manager develops the plan to use resources and resolve conflicts. The organizer develops a plan to use time, money and resources such that the ideas created will work.

- Phase 4: Completion. This phase covers the analysis of alternative methods, the decision of the plan of action and the execution of the task. Alternative methods to implement the task must be considered. Evaluator and Finisher team roles are important in this phase. The evaluator makes judgments on situations, plans, results and alternatives. The finisher follows plans and attends to the completion of the task.

Mumma considers that sometimes teams fail at reaching their goals because people only carry out the tasks they like omitting some of the phases that are essential for the team progress. For this reason, it is important for a team to be composed of people having different team roles. Together with the team roles theory, Mumma developed a test, through which people can know which their favorite work phases are and they can learn to cultivate the abilities to complement the absence of a certain role in some critical phases. Mumma's questionnaire is composed of 18 items, each of them associated to 4 hypothetical situations about how a person behaves in a group. The person that answers the questions has to score each situation in a 1-5 scale (without the 4). The aim of our work is to create an alternative way to make this detection without completing this extensive questionnaire.

2.2. Related Work on Team Role Detection

In the last years, with the advances in virtual organizations and the development of groupware systems, the concepts studied in the area of Psychology for the analysis and formation of groups have been implemented in Computer Supported Collaborative Work (CSCW) and also in CSCL. We can find in the literature articles that show that the group performance is better in those groups in which there is a balance between the social and task resolution abilities. For example, in (Costaguta et al., 2012) the authors present a multi-agent model that monitors students' participation in a group and proposes corrective actions when the group behavior is far from the ideal balance of team roles. In (Aritzeta et al. 2007) the authors reaffirmed the theory about balanced teams performing better over the long run than imbalanced ones. Similarly, in (Winter 2004) it is proposed a group model that combines Belbin's team roles with team members' knowledge, type of task that needs to be done, and group contextual factor (such as localization and time pressures). Other investigations also propose similar models applied to the Software Engineering Area in order to make teams with better performance and efficiency (Dafoulas 2001). In (Aguilar et al. 2007) it is proposed an intelligent agent that acts like another team member and takes the needed role in order to balance the team. Moreover, in (Licorish et al. 2009) a tool that forms groups using people information in order to avoid conflicts is proposed. This information comes from personalities, skills and weakness of group members.

As it can be observed in the works mentioned before, there are several researches that study group and team formation in virtual domains. However, there are few works that try to detect team roles in an automatic way. Some works in this direction are (Ou et al. 2005, Wang et al. 2002, Zancanaro et al. 2006). In (Ou et al. 2005, Wang et al. 2002) the authors proposed the use of Text Mining techniques to detect their students' member-roles. On the other hand, in (Zancanaro et al. 2006) a machine learning technique that detects functional roles automatically is presented.

In this work, we propose a novel technique for the automatic detection of team roles that is carried out in a CSCL environment and users do not notice the underlying detection process. In this way, students do not have to respond tens of questions, and we avoid the risk of students making mistakes because of wrong question interpretation or because of selecting random answers to try to finish the answering questionnaire as soon as possible.

3. PROPOSED TECHNIQUE

Our proposed technique is summarized in Figure 1. First, we observe users' behavior when they work using CSCL systems in order to register their interactions, which are actions users carry out to reach the goals of the team. Then, the information gathered is used by a Bayesian Network (BN) that enables the detection of the user's team role.

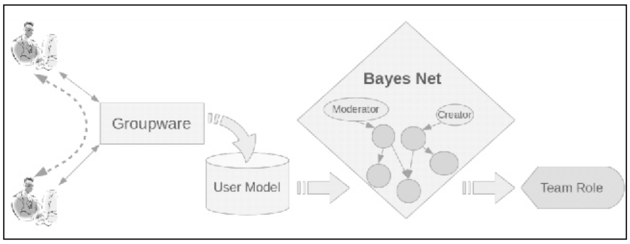

Figure 1. Overview of our approach.

3.1. Observed Interactions

We used a CSCL application developed at UNCPBA, Argentina, which captures students' interactions when they work collaboratively to reach a common goal. This application integrates different collaboration tools such as chat, discussion forums, calendars, instant messaging, among others. These components enable team collaboration, coordination and communication. The automatic detection of team roles is based on the analysis of the interactions in such a system. Particularly, we are interested in actions such as participation in debates, contributions to problem resolution, organization of activities, resource management, and support to team members, among others. These interactions are vital to determine team roles since the characteristics of Mumma's team roles can be associated with the actions and behaviors of people in a virtual team. Table 1 summarizes the association we propose between Mumma's roles, their features and the actions we observe in the system.

Role	Actions observed
Leader	Starts discussion topic s / Adds news / Assigns tasks / Clarifies goals
Organizer	Does a lot of tasks (efficient) / Finishes each task in a timely manner / Turns ideas into practical action
Manager	Accesses frequently to forum page / Takes part in discussions / Challenges team
Creator	Posts very few messages / Sends very few mails / Solves tasks / Proposes new ideas and solutions
Innovator	Has a lot of contacts in his/her book / Accesses frequently to forum page / Accesses frequently to the groupware at the beginning of the project / Sends negotiation messages
Moderator	Participates frequently in all discussions / Averts friction / Reaches agreement
Finisher	Finishes each task on time / Searches out errors
Evaluator	Analyzes options

Table 1. Actions observed.

In our application, most actions are extracted from non-structured textual communication channels such as chats or forums. Since analyzing these texts is a very difficult computational task, we decided to base our technique, at least initially, on structured CSCL systems. To obtain a structured collaboration system we propose the utilization of a set of opener sentences (Pohl et al. 1995). These opener sentences are phrases that can be selected by users when they want to post something on a forum or talk through a chat, and give a general idea of the type of contribution made. In this way, users who want to interact with their team members through a discussion forum will have to select first the opener sentence that best suits their collaboration. Below, we describe examples of the relation between actions and opener sentences used in our application.

For the leader:
- Action “clarify goals”: “To sum up...” / “So...” / “In other words ...”

For the organizer:

- Action “Turn ideas into practical action”: “We have to do...”

- Action “Efficient to solve tasks”: “What do we do now?”

For the manager:

- Action “Challenges team”: ““I can't understand. Could somebody ...?” / “Why....?”

- Action “Discuss”: ““But it could happen that...”

- Action “Behaves aggressively”: “I don't agree ...” / “I am not sure ...”

For the creator:

- Action “Proposes new ideas and solutions”: “Try...” / “I think we should try...” / “I think that... because...”

For the innovator:

- Action “Negotiation messages”: “What do you think? We try... or...” / “Please, somebody show me...” / “We should ask to...”

For the moderator:

- Action “Averts friction and discussion”: “Ok, I agree...” / “It's OK to me” / “Let's go on! We are right!” / “Sorry, but...”

- Action “Reaches agreement”: “All your opinions are right...” / “This goes well. Let's go on!” / “Do you agree...?” / “Go on...”

For the finisher:

- Action “Searches out errors”: “What do we need to consider?”

For the evaluator:

- Action “Analyses options”: “Instead of ... we should test...” / “Instead of that, we could...” / “Let us to suppose…”

3.2. Modeling Student's Behavior With A BN

A BN is a compact, expressive representation of uncertain relationships among parameters in a domain. A BN is a directed acyclic graph where nodes represent random variables and arcs represent probabilistic correlation between variables (Jensen 1996). The absence of edges in a BN denotes statements of independence. A BN encodes the following statement of independence about each random variable: a variable is independent of its non-descendants in the network given the state of its parents (Pearl 1988). A BN also represents a particular probability distribution, the joint distribution over all the variables represented by nodes in the graph. This distribution is specified by a set of conditional probability tables (CPT). Each node in the network has an associated CPT that specifies this quantitative probability information. A CPT indicates the probability of each possible state of the node given each possible combination of states of its parents. For nodes without parents, probabilities are not conditioned on other nodes; these are called the prior probabilities of these variables. In our problem, random variables represent the different team roles of Mumma's theory and the factors that determine each of the team roles. These factors are extracted from the interactions between the student and the collaborative system. Thus, a BN models the relationships between the team roles and the factors determining them.

To build a BN, it is necessary to complete three steps: identifying nodes and their possible values, establishing relationships between these nodes, and completing all probability tables. To identify nodes, we took each team role definition made by Mumma and extracted actions we could recognize through our groupware system. These actions are the interactions observed from users explained in Section 3.1; other more abstract actions, which were identified to make the understanding of the BN easier, and the eight team roles defined by Mumma. Nodes can get different values according to their type. For example, if a node represents an action that is identified through an opener sentence its possible values are “Yes” and “No”, which means the fact that a user has used that opener sentence in some post or not, respectively. Below we describe all the node types in our BN and their possible values. To establish the relationship between these nodes, we also took each team role definition and made relationships between those nodes which were related.

Nodes of type A represent an action observed through our groupware system. Some examples are “start a discussion topic” or “add news”. Possible values are “few”, “medium” and “many”. Nodes of type B represent an action which is identified through an opener sentence, such as “Reaches agreement”, “Clarifies goals”. These nodes can take the value “yes” or “no”. Nodes of type C represent an abstract action that groups specific actions. Examples of these nodes are “Avoid problems”, “Communication”, and they can take values “few” or “many”. Finally, nodes of type D represent a team role. The values might be “yes” or “no”.

Initially, probability values for independent nodes are assigned equal values. Then, the values are updated as the system gathers information about the student behavior. The probability functions attached to the independent nodes are adjusted to represent the new observations or experiences. Consequently, the Bayesian model is continuously updated while new information about the student's interaction with the system is obtained. On the other hand, the probability values contained in the different CPT were obtained via a combination of expert knowledge and experimental results. The expert knowledge was obtained from Mumma's definition of team roles. We took into account the influence of the different factors analyzed on the dimensions of the different roles. To determine the values experimentally, we asked a set of 50 Computer Science Engineering students to complete Mumma's questionnaire. Then, we let these students use the CSCL system and recorded their interactions with it. Information about their recorded behavior was used to determine the conditional parameters of the BN, in combination with the expert knowledge. It is worth noting that the values in the CPT are equal for all students, as they represent the strength of relationships between different behaviors and team roles. However, the values corresponding to independent nodes are different for different students as they represent the actions taken by a particular student and are obtained from each student's log file. Part of our BN is shown in Figure 2.

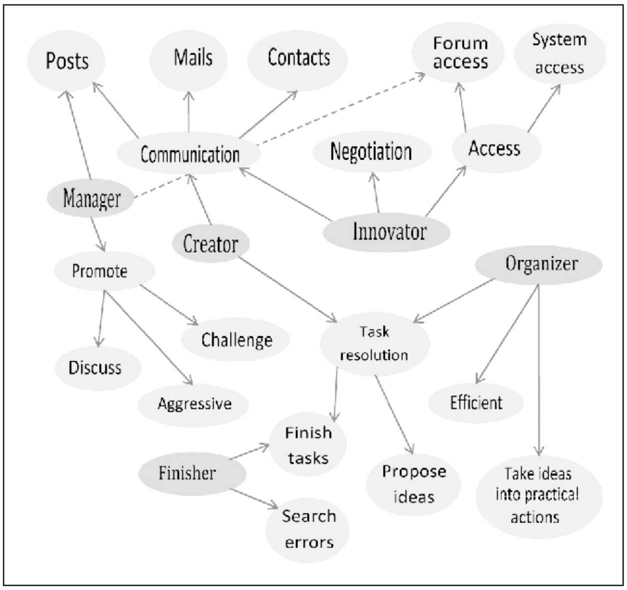

Figure 2. Part of our Bayesian network to detect team roles.

Once we have built the BN and we have determined the values of the probability tables, our goal is inferring the values of the nodes corresponding to the dimensions of a team role given evidences of the student's behavior with the system. Thus, we obtain the probability values of the team role node given the values of independent nodes. The role of the student is the one having the greatest probability value.

4. EXPERIMENTAL RESULTS

We used a data set obtained from the interaction of 94 students with and through a collaborative system. These students belong to 3rd year of the System Engineering career at our university. They took, during the 2nd semester of 2009, a course in which they use our groupware system with the objective of solving a series of exercises in groups. The population of students consisted of 81 males and 13 females. The students completed Mumma's questionnaire to determine their roles, which were considered as correct for our experiment. Figure 3 shows the team role distribution. We can see that our data set is not well distributed since only 2 students (2%) got the team role finisher over 24 students (26%) who got the leader one.

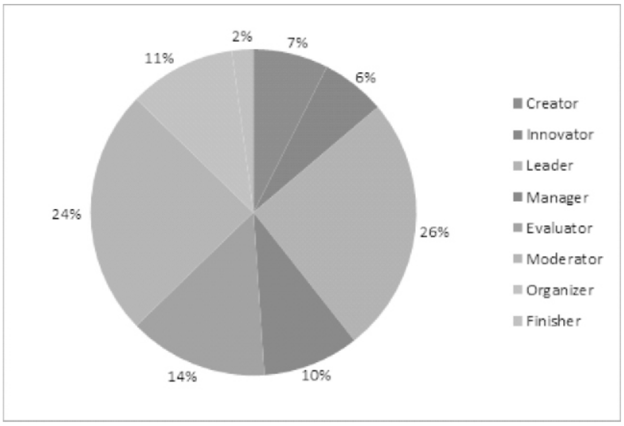

Figure 3. Team role distribution.

To evaluate our BN's detection, we set as evidences all leaf nodes (A and B type nodes) using each attribute in the instance. Thus, applying the Bayes' theorem and

Bayesian inference mechanisms we can know the probability of the remainder nodes (C and D type nodes). Then, we ordered D type nodes by their positive probability making that the most probable team role would be in the first position.

In order to evaluate our BN precision to detect user's team role we used two methods. First, we used a function that determines whether the detection is a hit or not. We consider a hit when one of the most three probable team roles is the same to the team role that is in the data set (real team role). Second, we used fitness function that gives points depending on the position in which the real team role is inside the ordered team role list detected by our BN (P(R)). Equation 1 shows how the fitness function works. The function f returns 1, 0.75 or 0.5 points if the team role R is in the first, second or third position in that list, respectively.

$$f(R) = \begin{cases} 1 & P(R) = 1 \\ 0.75 & P(R) = 2 \\ 0.5 & P(R) = 3 \\ 0 & otherwise \end{cases}$$

Figure 4 shows the results obtained. It is interesting to note that more than 56% of the tests obtained the first position and 76 instances were hits (81%). Our BN got f (R) = 67.00 and precision = 71%. We consider that these results are quite promising and hence, our technique is a viable alternative to determine team roles automatically in CSCL.

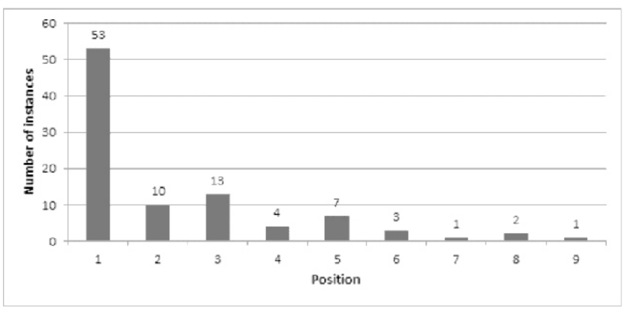

Figure 4. Precision of our proposed technique.

5. CONCLUSIONS AND FUTURE WORK

We have presented a novel technique for the automatic detection of students' team roles in groupware systems. We applied an Artificial Intelligence technique to the data obtained from the observation of students' actions in a CSCL system. The results obtained show that our proposal is a viable alternative to questionnaires, which are the most common tool to determine roles. Thus far, we have not found in the literature a similar approach for team role detection in CSCL.

With our approach, teachers can use an automatic technique to detect student's roles and use this information to build learning and working groups considering balance conditions in the roles involved to achieve better performance in their students.

Currently, we are working on text mining techniques for the detection of user's intentions when they contribute to the group. In this way we will not have to depend on opener sentences and structured tools. We are also planning to carry out more experiments with a higher number of users in order to validate the results obtained thus far.

ACKNOWLEDGMENTS

This work has been partially funded by ANPCyT PICT-2011-0366 and CONICET PIP No 114-200901-00381.

REFERENCES

Aguilar, R., de Antonio, A., & Imbert, R. (2007) Searching Pancho's soul: An intelligent virtual agent for human teams. In Proceedings of CERMA'07, 568–571, Washington, DC, USA.

Aritzeta, A., Swailes, S., & Senior, B. (2007) Belbin's team role model: Development, validity and applications for team building. Journal of Management Studies, 44 (1): 96–11.

Bales, R., & Strodtbeck, F. (1951). Phases in group problem-solving. Journal of Abnormal and Social Psychology, 46 (4): 485–495.

Belbin, R. (1981). Management teams: Why they succeed or fail. Oxford: Butterwoth-Heinemann.

Belbin, R. (1993) Team roles at work. Oxford: Butterwoth-Heinemann.

Costaguta, R, Schiaffino, S., & Fares, R. (2012) Fostering team-role balance in Computer Supported Collaborative Learning, New Horizons in Creative Open Software, Multimedia, Human Factors and Software Engineering, (pp. 82–90). Blue Herons.

da Silva, F. & Cesar, A. (2009). An experimental research on the relationships between preferences for technical activities and behavioural profile in software development. In Proceedings of the SBES '09, (pp. 126–135).

Dafoulas, G. & Macaulay, L. (2001). Facilitating group formation and role allocation in software engineering groups. In Proceedings of the ACS/IEEE International Conference on Computer Systems and Applications (AICCSA '01), (pp. 352–359), Beirut, Libano.

Jensen, F. (1996) Introduction to Bayesian Networks. Springer-Verlag.

Licorish, S., Philpott, A., & MacDonell, S. G. (2009) Supporting agile team composition: A prototype tool for identifying personality (in)compatibilities. In Proceedings of the 2009 ICSE Workshop CHASE '09, (pp. 66–73), Vancouver, Canada.

Mumma, F. (1992) Team-Work & Team-Roles. King of Prussia: HRDQ.

Mumma, F. (1994) What makes your team tick? HRDQ.

Ou, K., Wang, C. & Chen, G. (2005) Identify group roles by text mining on group discussion in a Web-based learning system. In Proceedings of 2005 International Conference Machine Learning and Cybernetics (ICMLC'2005), (pp. 5566–5572), Guangzhou, China

Pearl, J. (1988) Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers Inc..

Pohl, W., Kobsa, A., & Kutter, O. (1995) User model acquisition heuristics based on dialogue acts. In International Conference on the Design of Cooperative Systems, (pp. 471–486).

Wang, C., Ou, K., Liu, C., Liu, B., & Chen, G. (2002). Using a text miner to identify group roles of students in a Web-based learning system. International Journal of Learning, 9:1033–1043.

Winter, M. F. (2004) Developing a group model for student software engineer- ing teams. PhD thesis.

Zancanaro, M., Lepri, B. & Pianesi, F. (2006) Automatic detection of group functional roles in face to face interactions. In Proceedings of the 8th International Conference on Multimodal Interfaces (ICMI '06), pages