Evaluation of communication in web-supported learning communities – an analysis with triangulation research design

Andreas Harrer*, and Sam Zeini

Collide Research Group
University of Duisburg-Essen
Faculty of Engineering
47048 Duisburg, Germany
E-mail: harrer@collide.info
E-mail: zeini@collide.info
*Corresponding author

Niels Pinkwart

HCI Institute
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh PA, USA
E-mail: nielsp@cs.cmu.edu

Abstract: In current educational practice, web-based environments are an established means to support learning scenarios. In the study presented in this article, we use them as supporting tools for presence-learning scenarios and investigate the effects of electronic communication forms in this environment. Following the idea of triangulation, we utilised qualitative methods, statistical analysis and Social Network Analysis. Qualitative methods were used to classify the usage types of communication in Wiki and discussion forums. Based on these categories, we selected project groups with maximal variance of their communication behaviour and conducted Social Network Analysis to explore communication structures in detail. This research design is supported by the fact that no single communication form proved to be superior: combined usage produced better results with respect to the final scores. We concluded that more aspects of the respective communication forms should be taken into account to be able to compare them properly.

Keywords: web-based communication; presence-learning scenario; community portal; evaluation; triangulation design; Social Network Analysis (SNA).

Reference to this paper should be made as follows: Harrer, A., Zeini, S. and Pinkwart, N. (2006) 'Evaluation of communication in web-supported learning communities – an analysis with triangulation research design', *Int. J. Web Based Communities*, Vol. 2, No. 4, pp.428–446.

Biographical notes: Dr. Andreas Harrer is Research and Teaching Assistant at Universität Duisburg-Essen. He received his PhD in Computer Science from the Technical University of Munich with a PhD thesis in the area of intelligent support of collaborative learning interactions (Computer Supported Collaborative Learning or CSCL). His current main research interests are

the analysis and intelligent support of group learning as well as software engineering principles for educational systems. He is Speaker for the special interest groups 'Artificial Intelligence and Education' of the 'Kaleidoscope' Network of Excellence and of the 'Intelligent Teaching and Learning Systems' group within the German Society of Computer Science (GI).

Dipl. Soz.-Wiss. Sam Zeini is Research Assistant at the Universität Duisburg-Essen. He received his diploma in Social Sciences from the Universität Duisburg-Essen with a diploma thesis about the pragmatics of knowledge in organisational contexts. He is currently working on the VIP-NET Project funded by the German Federal Ministry of Education and Research (Project No. 01HU0128) and conducted by the COLLIDE research group. His research interests are empirical methods and social aspects of Computer Supported Cooperative Work and CSCL.

Dr. Niels Pinkwart graduated as a High School Teacher in Mathematics and Computer Science from Duisburg University in 2000, and received his PhD in Computer Science from the University of Duisburg-Essen in 2005. He is currently a Post-Doctoral Fellow at the Human Computer Interaction Institute of Carnegie Mellon University, working on a project about supporting law students in the acquisition of legal argumentation skills. His research investigates interactive and collaborative technologies with a specific focus on applications in the field of learning – the design of distributed and collaborative software systems which support groups of learners intelligently in their tasks.

1 Introduction

In current educational practice, web-based environments are established tools to accompany learning scenarios. In contrast to other computer-based support methods, web-based tools have some inherent practical advantages: they normally do not require the user to install any software, and a significant number of learners nowadays already have some experience in browsing the internet and are thus used to the underlying usage patterns.

There are, indeed, countless variants of how the internet is used in learning contexts. A traditional (and in practice frequently occurring) usage mode is that web pages serve as a more or less static information source, *e.g.*, to provide the students with resources they need for a course. Other established forms are web-based intelligent tutoring systems (*e.g.*, Kay and Kummerfeld, 1994; Brusilovsky *et al.*, 1996; Kinshuk *et al.*, 2003). As such, the main purpose of the system is to teach rather than to be a learning resource. A number of web-based environments like Future Learning Environment (FLE3) (Dolonen *et al.*, 2003) take into account social perspectives of learning, and offer means of communication or collaboration as a central element of the web-based learning support. This communication/collaboration support can take several forms, from 'simple' discussion forums to constructivist-oriented tools like Sistemas de Representación (SDR) (Madrazo, 2003) or KnowledgeForum, which emerged as a web-based commercial version of the well-known Computer-Supported Intentional Learning Environments (CSILE) (Scardamalia, 2004).

In addition to these observable differences in the function of existing web-based environments for learning support, the usage context of these systems also varies considerably: some approaches are related to distance-learning scenarios (which of course is an obvious application area for web-based tools), while others focus on the support of presence courses. Another dimension that characterises the usage context is the learning group size, which can vary between very small groups (or even isolated single users that do not interact with others) and large communities with their special needs (Kim, 2000; Gaudioso and Boticario, 2003). Further distinguishing criteria are, for example, the course type (*e.g.*, lecture vs. seminar) or the age of the students.

Obviously, the educational context does have an impact on the suitability of specific web-based tools. The group size, for example, is an important factor that determines the potential for interaction in the learner community. Another example for the mentioned impact is that in distance-learning situations, the computer support is often used to make up for the lack of personal (face-to-face) interaction and communication. Therefore, distance-learning tools can often be characterised as learning *environments* in a narrower sense, whereas environments for presence-learning situations are often *parts* or *tools* used in learning environments.

The support of presence lectures at the university level through web-based environments as we find it today is a typical intermediate case between presence- and distance-learning situations, sometimes denoted by the term 'blended learning' (Sauter *et al.*, 2004): often, the lecture is done physically, but a lot of supporting actions are delegated to a web-based environment owing to lack of time or of university staff, or other constraints. It is obvious that there are a number of functions of the web-based support environment that would be convenient for students and teachers in this case (*e.g.*, communication facilities, resource collections, the management of exercises if appropriate). Of course, a central important question here is whether this web-based support really does improve the learning results of the students. Web-based support for presence lectures meets most of the characteristics of Online Knowledge Communities (*e.g.*, goal of developing knowledge, continuous interaction between the participants, highly complex and unstable context), so that the questions of how to design a suitable meeting place and how to enable and support continuous interaction among the participants seem to play a key role for success (de Vries and Kommers, 2004).

Pinkwart *et al.* (2005) presented the internet Portal to Augment Learning (iPAL) web portal and its use in a blended-learning scenario of the described characteristics, a mixture of presence-learning parts and web-based parts. The paper shows some investigations that analyse the interrelations between active usage of the forum embedded in iPAL and the student's final mark in the examination. Indeed, a positive correlation was observable, and one of the particular results in that study was that an above-average system usage is very likely to go along with a good mark.

However, interesting questions remain. In order to really understand the relations between the usages of the web-based environment in the learning group and the learning outcome, the applied methods (pure counting of postings and correlation to marks) might be too superficial. More in-depth studies, which concentrate on the collaboration processes in the environment and their influence on the learning outcomes, are required.

Here, the general problem of evaluating online communities interacting through web portals comes into play: according to Preece *et al.* (2004), most standard evaluation methods fail in accounting for important social factors. Preece *et al*. propose ethnographic or heuristic techniques to make up for this. In contrast to this, our approach presented in this article makes use of a method mix which includes Social Network Analysis (SNA) as one of the components. This allows for including social dimensions in the evaluation (at least at the level of hypothesis generation), and in particular serves the purpose of detecting and identifying roles in the community well, which is an essential part of trying to understand group structures (Madanmohan and Navelkar, 2004). We have taken the encouraging results of the study carried out by Pinkwart *et al*. (2005) as motivation and redesigned the iPAL web portal to be able to conduct a finer analysis of the system usage and its comparison to the learning outcomes. In the following, this article first briefly describes the basic features of the iPAL portal and the target scenario in which we have applied it, and then presents the evaluation results together with our interpretation.

2 The setting: lecture and web-based communication environment

The course that we investigate in this article is the lecture 'Software Engineering' held in the summer term 2004 at the University of Duisburg-Essen in Germany. The course is taken mainly by second-year students of Computer Science. It consists of a lecture with accompanying exercises as the presence-learning scenario. Additionally, we augmented this with remote support by providing a web portal with integrated communication facilities.

As the technical starting point, we took the iPAL portal described in Pinkwart *et al*. (2005). This software package uses the royalty-free platform Postnuke for the basic features of a web portal, such as user management, polls and news. To stimulate communication and exchange we added freely available extensions, such as a discussion forum and an upload/download module. To integrate the presence support and web-based support with each other, we developed a module to upload the electronically captured lecture notes (Bollen *et al*., 2003) to the portal. In the newest version, this can be done automatically direct from the lecture hall. We also implemented a management module for the (presence) exercise groups and the electronic submission and grading of exercises. This environment had been used in the previous course and was described and discussed with preliminary results in Pinkwart *et al*. (2005).

For this term's lecture, we used feedback of the learners and our own observations (while evaluating the last portal version) to enhance and refine the web-based lecture support environment. On one side, we improved the integration of the lecture notes in electronic format, where the students can now directly attach annotations to the notes or link specific notes indirectly to the discussion forum for general discussion of the topic. On the other hand, we found out that a regular discussion forum with flat thread structure cannot represent the structure of the communication flow properly, because information on who is responding to whom gets lost in the flat thread. Because we wanted to explore the communication structure in detail using techniques of SNA (Wassermann and Faust, 1994), we extended the discussion forum module according to our requirements: the new discussion forum represents tree-structured discussion threads (*i.e.*, postings are made

referring to one specific posting, not to a general topic like in flat threads) internally and thus keeps all the information for detailed communication analysis. From the user interface it preserves the usual look and feel of a regular discussion forum, *i.e.*, flat thread enhanced by links to the referenced posting.

The exams were conducted through a mix of a small group (three to five students) project of four weeks duration and oral exams taken after the project submission. To support the small project groups with proper communication infrastructure, each group was given a small group discussion forum, a Wiki and an access to a Concurrent Versioning System (CVS) server:

- The small group discussion forum was meant for communication within the project groups and with their assigned 'customers', our student tutors taking the role of customer of the software project to be developed.
- The Wiki[1] was introduced to the students in the lecture as a means of co-constructively editing and refining living documents, which can be used for finding common ground on specific terms by defining their interpretation (this met with great success in the Wikipedia on the web).
- The CVS server supported the distributed software development by taking responsibility for version management and conflict management in case of concurrent modifications of source codes and project documents.

In the following sections of the article, we will analyse and discuss the usefulness of these communication facilities and their impact on group structure and dynamics, as well as on the outcome of the exams. This is meant to shed light on our preliminary results (Pinkwart *et al*., 2005), which showed that a strong participation in the lecture's discussion forums correlated with the achieved grades. At this point we investigate more deeply the use of a variety of support tools for project work. Our hypothesis is that using computer-based communication infrastructure facilitates the success of project work: here, we put a specific focus on relating the different communication means with each other. The question of whether there is a key communication infrastructure crucial for success or if synergy/balance of different tools proves to be effective is especially a focus of our study.

3 Research design and methodology

The methodology of the study can be characterised as a mixed-method design, following the idea of triangulation (Denzin, 1989). The decision was to use qualitative methods, statistical analysis and SNA (Wassermann and Faust, 1994). This research design allows for using the results of one applied methodological approach as an interpretation context for the other methodological pathways.

- Qualitative analyses of the forums and the Wiki

  Qualitative methods are suitable for understanding new phenomena. In triangulation designs qualitative methods are usually used with the aim of building typologies and hypotheses. In our case the building of hypotheses was derived from the question of the differences between the typologies we found, and also by asking how these are affected by other factors.

- Social Network Analysis (SNA)

  SNA is intended for analysing structures. In contrast to quantitative methods which can analyse structures indirectly through the operationalised properties of the analysed cases, SNA allows the reconstruction of social structures, for example communication paths. In our study we used the typology derived by the qualitative analyses for sampling the most interesting groups, in the sense of having the highest variance in the way they organised their project. We assume that analysing these networks could point out some important aspects of their course of action.

- Statistics

  Based on categorisation of the groups with different types of Wiki and forum usage by qualitative analyses, statistical analysis serves to explain the differences between the groups. This was done by formulating hypotheses. In this case some indicators were outlined to test the hypotheses.

- Long-term statistical analysis

  Since iPAL was used to support a previous course, we decided to compare the results of both courses. In addition to the result comparisons, we were also able to make some long-term analysis, because 75% of the students from the current course were also present at the course.

The qualitative analysis was mainly done by long-term observations through the teaching staff and by analysing the content of the Wiki and the forums. The data for the SNA and the statistical analysis were extracted from the mySQL database used by iPAL and the log-files of the CVS. The dataset for the statistical analysis contains 73 cases that are represented by students who carried out all parts of the examinations. The dataset extracted from the CVS log-files contains 20 groups which were represented by the project groups that carried out a software project as part of their examinations.

4 Qualitative and quantitative results from Wiki and forum usage

To understand the usage of the Wiki and the forums within the different project groups, we analysed the content and the creation process through its versions, as well as the forums, qualitatively. We found out that the Wiki usage varied widely in separate dimensions:

- On the one hand, the *interactivity of the construction*, *i.e.*, number of different authors, number of versions and scope of changes between versions, varied: some groups made small and frequent updates/modifications, some had few but rather big changes between versions. Additionally, some Wiki pages seemed to have been the 'property/responsibility' of one person, because they were edited exclusively/mainly by one person.
- On the other hand, the *content* and thus the *purpose of the usage* varied: we found and indexed four categories of usage of the Wiki:

a Project management

The Wiki is used to coordinate the team members' activities and document their planning. Updates are usually done when replanning, rescheduling and making counterproposals. The final version is (probably) the documentation of the project process as it happened in reality.

b Clarification of terms/glossary construction

The Wiki is used to find a common ground and understanding of central terms and concepts for the project. Updates are usually done when introducing or defining new terms. The final version is a glossary of used concepts and terms of the project.

c Reference list

The Wiki is used as a common index to outside resources. Updates are usually done when giving new references and links.

d Coding conventions

The Wiki is used to produce a style guide for programming and/or documenting code. Updates are usually done when conventions are proposed, changed or retracted. The final version represents the conventions to be used within the project.

We analysed the interactivity of construction and the usage type of the different project groups. Of the 20 project groups, ten used the Wiki extensively, while ten used it very little or not at all. (Some groups exchanged ICQ numbers, e-mail addresses or mobile phone numbers, or worked with agile software engineering methods with team programming, so we assumed that communication happened mainly outside of our support environment in these groups.) For the ten groups using Wiki, we manually indexed the type of usage and got the results shown in Table 1.

Table 1 Categories of usage for the project groups' Wiki

Group	*Purpose of usage*	*Interactivity of construction*	*Average score*
A	Project management, glossary construction	Few versions, large differences	71.125
B	Project management	Frequent versions, small differences	68.667
C	Reference list	Frequent versions, small differences	63.0
D	Project management	Frequent versions, small differences	58.375
E	Project management	Frequent versions, small differences, some pages with exclusive responsibility	77.625
F	Coding conventions	Few versions, large differences	84.625
G	Project management	Few versions, large differences	65.5
H	Project management	Frequent versions, small differences	d.n.f.
I	Project management, reference list	Frequent versions, small differences	85.25
K	Glossary construction	Few versions, exclusive responsibility	75.8

Note: N = 73

As a result of the forum analysis, four different categories of usage can be distinguished:

1 The first category shows a very structured behaviour in using the forums. In these forums we usually found more threads than in other groups. The topics of the threads were structured, but threads were short.

2 The second category posted just a few but long threads.

3 The third category posted a great number of threads, and some of the threads were very long. In this case we also observed a differentiated topic structure.

4 The fourth category used the forums just for planning meeting dates.

Table 2 shows a summary of the forum usage. It is statistically significant that Category 3 has the greatest average number of postings and also the best results with respect to the average score. This category also produced the greatest number of files (383) in the CVS and the second-most numerous versions after Category 1. Another interesting result can be seen in Category 2. One of the project groups within Category 2 decided to use the agile programming paradigm and another project group chose a modular approach based on division of labour. They were put under a subcategory, which obtained an average score of 86 while the other two project groups within Category 2 which had not specified any programming approach obtained just an average score of 53 points. This bias has to be mentioned because the agile approach usually shows an extensive face-to-face communication structure and the modular approach shows a rationalised communication structure in favour of the division of labour concept.

Table 2 Categories of usage for the project groups' forums

Category	*Forum structure*	*Average postings*	*Average score*
1	Differentiated topics structure, short threads	11.1333	75.2667
2	Long threads, few topics	5.82352	70.5882
3	Differentiated topics structure, long threads	16.5714	87.4286
4	Meeting organisation	7.25	66.8750
5	Mixed	4.75	58.3750

Note: $N = 73$

Counting both concepts together shows that the project group with the highest average score (93 points, $T = -6.29$, significance at $P < 0.001$) belonged to Category 1 of the forum characterisation and showed no extensive Wiki usage at all. All members of this group were also present at the course we analysed last term. The group with the lowest average score (40 points, $T = 3.51$, significance at $p = 0.001$) used the forum but not the Wiki. It is also interesting that there is no significant difference in the average scores when classifying the project groups into categories which just used Wiki or forum, or used both Wiki and forum.

5 Social network analysis

For our plans to investigate the patterns of usage of the discussion forum and the resulting communication structures, we followed SNA, a well-established method for the exploration of social structures through mathematical techniques (Wassermann and Faust, 1994). SNA is based on graph theoretical ideas (relating the actors to each other in a matrix/graph representation) and provides an instrument to compute precisely traits of social networks, such as density of a network, centrality and prestige of actors, or centralisation of the whole network (Wassermann and Faust, 1994). It has already been applied in distance-learning scenarios (Reffay and Chanier, 2003) as a means to measure the cohesion of collaborative learning groups. In our scenario of mixed modes of presence teaching and web-based support, we wanted to find out if typical communication structures would evolve in the community of students and staff within the web-based support system and if there was any relation between communication structures, project quality and final results of the students.

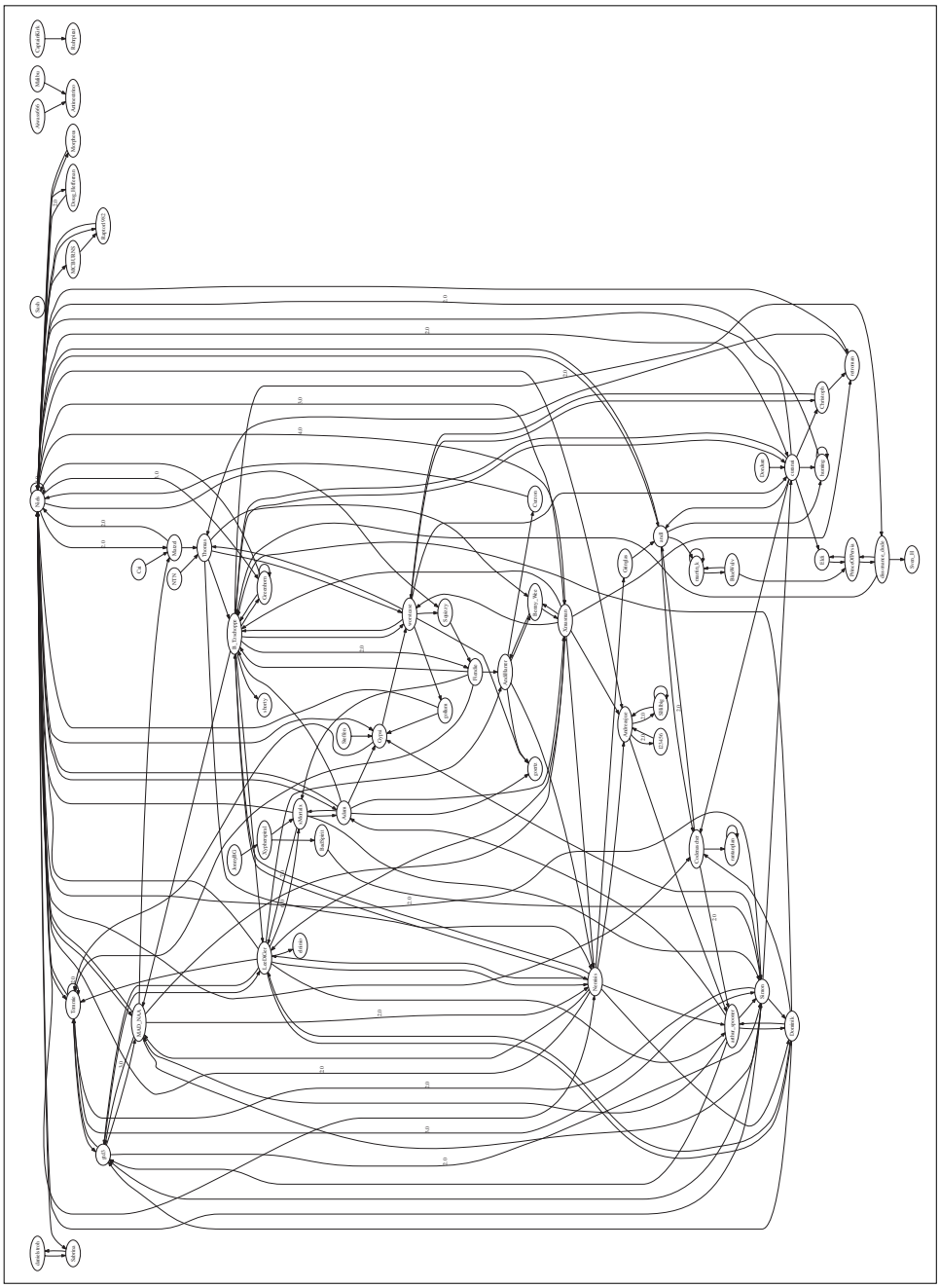
A relevant question was how to define a 'communication act' within our iPAL system. We chose to consider 'direct active communication' in this study, which manifests itself as a direct *answer* of an actor to another actor's posting in the discussion forum. Other types of relations between actors like 'interest in', when an actor *reads* another actor's posting, will be the subject of future work. To gain information about direct communication, we had to modify the standard discussion forum so that the exact reference of each posting (*i.e.*, the other posting to which a posting was created as a reply) was represented within the system (cf. Section 2). In previous work we already developed a tool for SNA (Harrer, 2004a) for a specific semi-structured conversation forum (Harrer, 2004b). To reuse this tool with our iPAL system, we transformed the postings' database representation (on mySQL basis) into an XML representation that could be processed by our SNA tool.

For detailed analysis of the communication structures, we opened the general discussion forum to all users of the iPAL system, and additionally provided separate forums for each project group and their 'customer' (cf. Section 2). Owing to article length restrictions, we will focus on selected SNA features which are applied to the general forum and contrast/relate it with a few project groups with distinctive project processes, communication structures and project results. For visualisation of the communication structures, we used network graphs (similar to sociograms (Moreno, 1951)) that have been produced automatically by our SNA tool using the graph layout software dot.[2] Among the SNA traits we present here are the centrality of one actor, the centralisation of the respective network, and the prestige of an actor, all of them computed based on the actors degree within the graph (our tool also provides betweenness and proximity-based computation).

The general discussion forum which is presented in Figure 1 had 64 persons creating 276 postings. The computed value for degree-based centralisation is $C_D = 0.283$ (0 stands for a completely balanced network where every actor has the same centrality, 1 for a completely centralised network (star topology)). This shows that the network had some 'key players' (including one central person from our staff and some very active students), but also that, in general, the network was not dominated too much and was used by most students to communicate among themselves as well as with the staff.

This is shown in the average of individual actors' centrality avg($C_D(n)$) = 0.044, which means that the general centrality of actors was quite low, so that nobody could be called 'hub' in this network.

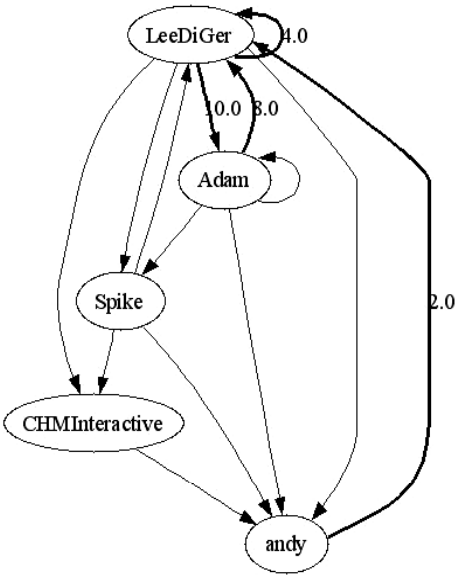
Figure 1 Sociogram of the whole posting community



Regarding the project groups which typically consisted of three to five students and one 'customer' (staff playing the role of the groups' customer), we were mainly interested in differences between the specific groups (because of the much smaller size of the resulting networks and their close cooperation during the project time of four weeks, a direct comparison between the general network and the project group networks would be speculative at best) and relations between communication structure, project organisation and final outcomes. Driven by the qualitative analysis, the resulting typology and the concept of maximal variance, we present three selected groups which are distinctive with respect to the way they communicated with each other and with their customer, and their general project organisation:

Group 1 (no Wiki usage, little CVS, long threads with few topics), shown in Figure 2, had a relatively high centralisation $C_D = 0.5$ of the network (considering the small number of actors), with one student as central actor (centrality $C_D(p) = 1.0$ and prestige $P_D = 0.75$), the customer (Adam) with an amazingly small prestige of $P_D(c) = 0.25$ (in fact the smallest in this network), and the other actors with centrality ranging from 0.25 to 0.75, prestige from 0.5 to 1.0. This group indeed had problems with internal communication (inside and outside the iPAL system), which led to limited involvement of their customer, separate development of project subparts and massive integration problems for the project submission. The consequence was a much worse project outcome than the individual skills of the group members would suggest.

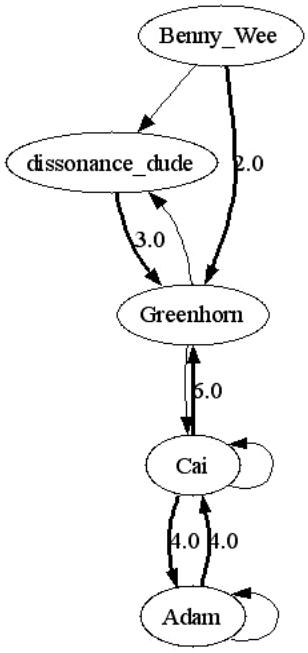
Figure 2 Sociogram of Group 1



Group 2 (no Wiki use, highest CVS, differentiated topics with short threads), shown in the sparsely connected graph in Figure 3, had a small centralisation $C_D = 0.125$ of the network and a quite low individual centrality $C_D(n)$ of the members, ranging from 0.25 to 0.5. The prestige $P_D(n)$ varied from 0 to 0.75, with the customer having 0.25 for both

centrality and prestige. This is explained by the specific process and distribution of labour this group chose: one of the members (Cai) was assigned to be 'the Key Account Manager' and exclusively communicated with the customer (Adam), both in the forum and in personal meetings. Since the division of labour and the whole project plan was followed, the project outcome resulted in the highest score among all the project groups. This group used other support facilities that we provided extensively, especially the CVS versioning system with more than 140 files and 1400 versions.

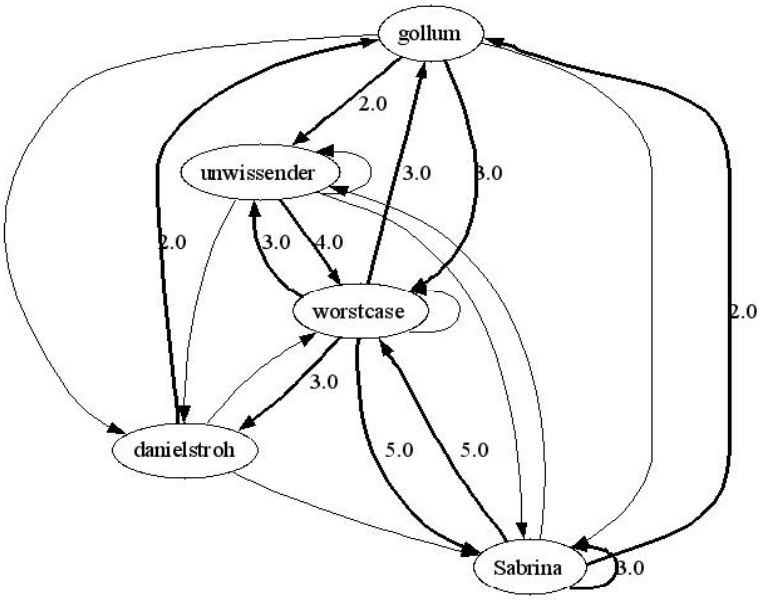
Figure 3 Sociogram of Group 2



Group 3 (extensive Wiki and CVS usage, differentiated topics with long threads), shown in the densely connected graph in Figure 4, also had a small centralisation $C_D = 0.1875$ but a consistently high individual centrality $C_D(n)$ ranging from 0.75 to 1.0 and prestige of $P_D(n)$ between 0.75 and 1.0. The customer (Sabrina) was intensively involved with centrality $C_D(c) = 0.75$ and prestige $P_D(n) = 1.0$. The use of all the support facilities we provided, *i.e.*, group forum and Wiki for project management and discussion, and CVS for code management, led to a well-coordinated project which scored second among all the project groups.

While reviewing the SNA results we found that the exclusively structural analysis might not be sufficient to explain the process and outcome of the group's work, but with the additional information we had as creators of the course most of the phenomena could be explained utilising both SNA and the process knowledge. Thus we think that especially in mixed presence/web-based scenarios, SNA can be a valuable help in understanding and interpreting communication structures.

Figure 4 Sociogram of Group 3



6 Statistical analysis

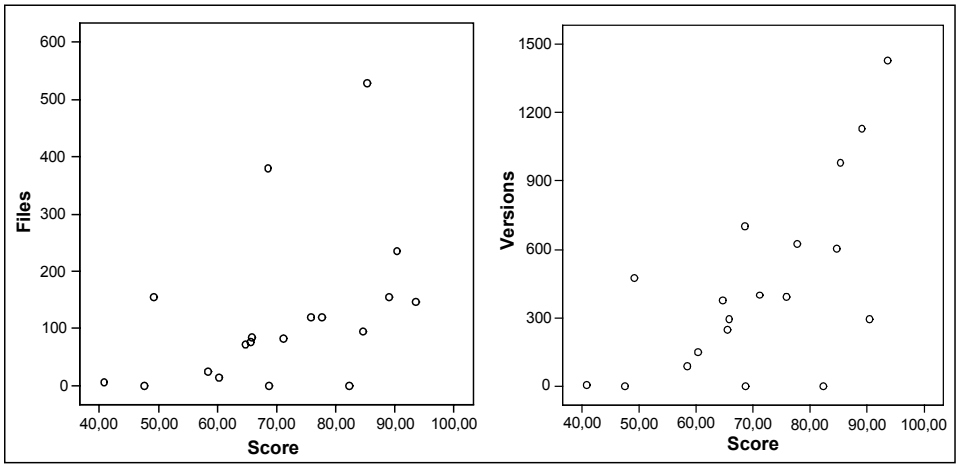
In addition to the statistical analyses that were made with respect to the qualitative analyses, in this section we will present two other statistical analyses in a global context.

6.1 CVS usage

The previous analyses aim to relate the behaviour of using the CVS system with the results of the course and also to the posting behaviour. For this purpose we aggregated the dataset we got from the teaching databases with data we could extract from our CVS system. Since the cases in the teaching datasets are represented by students but the CVS system could just be represented at the project group level for statistical analysis, we had to compute some new variables to be able to aggregate both data sources.

The dataset shown in Figure 5 represents 20 project groups and the average score, average number of postings, the number of files produced and the number of revisions of the files made by each group. The interesting outcomes are a) that there is a middle-strength correlation (0.541, significant at the 0.05 level, Spearman) between the number of files each group produced and the average score each group obtained in the course and b) a middle-strength correlation (0.571 significant at the 0.05 level, Spearman) between CVS revisions made by each group and the average score each group obtained.

Figure 5 Scatter plots of the CVS usage related to average scores

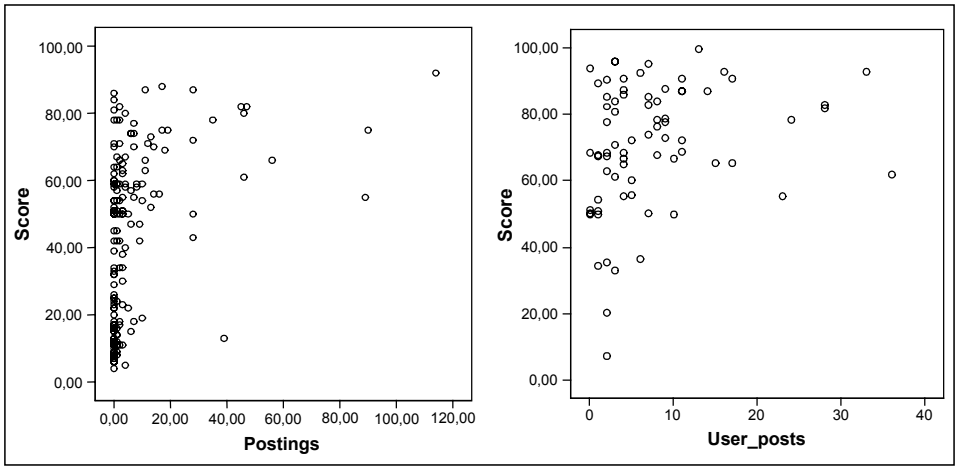


The other hypothesis we had was that there could also be a correlation between the average number of postings for each group and the CVS usage behaviour. In this case we could not find any correlations between the average number of posts and the number of files or versions that were produced using the CVS system.

6.2 Long-term analysis

As mentioned in Section 3, the approach presented in this article is particularly based upon an evaluation study that we carried out last term. Thus it is self-evident that we should compare the current results with the past evaluation. In the previous study (Pinkwart *et al.*, 2005), we were interested in finding out if there was a relationship between the number of posts from the students in the iPAL forums and the scores they reached in the course. We decided to reproduce the question in this study as a follow-up approach to compare results. Figure 6 shows the corresponding scatter plots.

Figure 6 Scatter plot of the relationship between postings and scores (left: past course, right current study; please note the different scales)



In this case we can observe that there was a stronger relationship between the number of postings and the average score (0.485, significant at the 0.001 level, Spearman) in the previous study than in the current study (0.320, significant at the 0.01 level, Spearman). This result led us to assume a fortification of personal relationships, and thus more direct (for us non-observable) communication between the students, since 75% of the students in the current course knew each other from the previous course. The hypothesis that this was caused by usage of Wiki could not be confirmed, since the students who used the Wiki did not show a significantly different posting behaviour (on the average) compared to the students who did not use the Wiki. This result led us to look at the differences between the 75% of the students ($N = 55$) who were present in both courses. In the current case the students received an average score of 71 points in comparison with the previous course, where they reached an average of 61 points. This difference is significant ($T = 4.72$, $p < 0.001$), and there is a correlation between the scores of the pairs (0.643, $p < 0.001$) that can be interpreted to mean that in most of the cases, students who received a high average in the previous course received a high average in the current course, too. On the other hand, we could not see a significant difference between the posting behaviours within the compared courses. In these cases the enhancement of the scores obtained by the students could also depend on the content of the courses.

7 Conclusions

In this article, we used mixed-method design to evaluate communication processes and structures within the web-based support system iPAL, which was used for a presence university course. Following the idea of triangulation, we utilised qualitative methods, statistical analysis and SNA. Qualitative methods were used to classify the usage types of the communication facilities Wiki and discussion forums. Based on these categories, we selected project groups with maximal variance in their communication behaviour and conducted SNA to explore communication structures in detail. The SNA of the whole learning community produced a non-centralised network, which complies with the large variety of communication facilities student subgroups used in the project work, according to their own choice. This degree of freedom was intended by the pedagogical approach. Indeed, this is supported by the fact that no single communication form proved to be superior. In fact the combined usage was shown to produce better results with respect to the final scores. This leads us to assume that more aspects of the respective communication forms should be taken into account to be able to compare them properly.

At this stage, SNA within a triangulation research design is relatively new to the CSCL research, but there are some comparable studies to refer to. For example, Cho *et al.* (2002) analysed a presence course (supported by a mailing list and a web-based discussion forum) by using statistics and SNA. Regarding Social Navigation (Höök *et al*., 2002), one of their main finding is that necessary URL's posted by central actors in the network received a higher peer response. They also found out that the mailing list generates more peer response than the discussion forum and described this phenomenon by the 'push' character of e-mail communication. This leads us to also integrate mailing list functionality within our iPAL system and integrate also into the analysis procedure. Another relevant result from similar studies also postulates taking weak ties within the network analysis into account (*e.g.*, Haythornthwaite, 1999). According to Whittaker *et al.* (1998), who stated that cross postings between different

groups increase interactivity, (weak) ties between groups within the iPAL should also be included in further evaluation studies. It should also be mentioned that some of the studies researching long distance scenarios (*e.g.*, Reffay and Chanier, 2003) point out the integration of synchronous communication forms in the evaluation process. Since iPAL was designed to support presence-learning situations, we are aware of the meaning of synchronous communication, which often takes place face-to-face. We do not intend to stress these situations by accounting for them in a quantifying analysis approach, but they can be taken into account by qualitative analysis, for example interviewing the teaching staff and the students.

8 Outlook

The issue of using mixed-method approaches for the analysis of complex learning scenarios has been taken up recently by different research groups, for example in Martínez Monés *et al*. (2003). Up to now this work has been done pragmatically according to the needs of the study at hand. Yet an integrated way of considering different methods for analysis has not been produced. An approach for a framework for analysis methods and their combinations has been started within the Kaleidoscope Network of Excellence in the sub-activity 'Interaction Analysis'. Besides the conceptual integration of nine European research groups' methods in a 'unified framework' (ICALTS, 2004), the discussion and definition of common formats to exchange analysis data and to make analysis components interoperable is a major topic in this initiative.

Our efforts at providing a tool suite for both learning support environments and analysis tools are along the same line of thought: to reduce the effort in creating mixed-method analyses in future scenarios, we aim at standard ways for setting up the learning support environments and automating the analysis methods that are appropriate for that setting. A first step that is finished now is the guided installation of whole web community portals with different communication facilities; given administrator rights and the availability of a mySQL database to store the portal information, the installation takes only approximately five minutes and requires little administrative expertise. A second step is the definition of an internal standard format for our analysis tools, so that the output produced by one tool can be used as input for additional analyses. With a versatile converter tool, the database content of discussion forums and other communication means, such as mailing lists, can be converted for use with SNA techniques. By providing this tool suite to the public audience, such as this journal's readership, we hope that we can contribute to a wider application of the evaluation of support environments for web-based communities.

We also plan to enrich the instruments available within our community portals for communication and collaboration. The above-mentioned inclusion of mailing lists has already been achieved at the analysis level, with SNA analyses possible on the mailbox archives. Besides the mainly asynchronous communication facilities embedded in community portals, we also aim at using collaborative shared workspace applications and analysing the collaborative processes therein: while we have some experience in isolated consideration of synchronous collaboration in small groups (McLaren *et al*., 2005; Harrer *et al*., 2005), the complex interrelations between synchronous small group and asynchronous community phenomena are largely unexplored and challenging.

A pilot study with our Greek research partners, who already conducted a large-scale distance experiment with the Hellenic Open University (Xenos *et al.*, 2004), is planned for the coming months. This will combine the use of the synchronous shared workspace tools Synergo (Avouris *et al.*, 2004) and Cool Modes (Pinkwart, 2003) in collaborating with mixed Greek-German dyads with discussion and asynchronous exchange with the whole student community in a web portal.

Acknowledgements

We thank our students Marcel Dauven and Oliver Lemm for their implementation work in Social Network Analysis, and Steffen Lohmann and Stefan Vetter for their support in the design of the iPAL system.

References

Avouris, N., Margaritis, M. and Komis, V. (2004) 'Modelling interaction during small-group synchronous problem-solving activities: the Synergo approach', *Proceedings of the Workshop on Designing Computational Models of Collaborative Learning Interaction at the 7th Conference on Intelligent Tutoring Systems (ITS)*, Berlin: Springer Verlag.

Bollen, L., Pinkwart, N., Kuhn, M. and Hoppe, H.U. (2003) 'Interaktives präsentieren und kooperatives modellieren. Szenarien akademischen lehrens und lernens in informatik und naturwissenschaften', in M. Kerres and B. Voß (Eds.) *Digitaler Campus. Vom Medienprojekt zum Nachhaltigen Medieneinsatz in der Hochschule*, Münster: Waxmann, pp.295–304.

Brusilovsky, P., Schwarz, E. and Weber, G. (1996) 'ELM-ART: an intelligent tutoring system on World Wide Web', in C. Frasson, G. Gauthier and A. Lesgold (Eds.) *Proceedings of the Third International Conference on Intelligent Tutoring Systems*, Berlin: Springer Verlag, pp.261–269.

Cho, H., Stefanone, M. and Gay, G. (2002) 'Social network analysis of information sharing networks in a CSCL community', in G. Stahl (Ed.) *Proceedings of the Conference on Computer Support for Collaborative Learning*, Mahwah, NJ: Lawrence Erlbaum, pp.43–50.

de Vries, S. and Kommers, P. (2004) 'Online knowledge communities: future trends and research issues', *International Journal of Web Based Communities*, Vol. 1, No. 1, pp.115–123.

Denzin, N. (1989) *The Research Act. A Theoretical Introduction to Sociological Methods*, Englewood Cliffs, NJ: Prentice Hall.

Dolonen, J., Chen, W. and Morch, A. (2003) 'Integrating software agents with FLE3', in B. Wasson, S. Ludvigsen and U. Hoppe (Eds.) *Proceedings of the Conference on Computer Support for Collaborative Learning*, Dordrecht: Kluwer Academic Publishers, pp.157–161.

Gaudioso, E. and Boticario, J. (2003) 'Towards web-based adaptive learning communities', in U. Hoppe, F. Verdejo and J. Kay (Eds.) *Proceedings of the 11th International Conference on Artificial Intelligence in Education*, Amsterdam: IOS Press, pp.237–244.

Harrer, A. (2004a) 'Analysis and intelligent support of learning communities in semi-structured discussion environments', in M. Bramer and V. Devedzic (Eds.) *Artificial Intelligence Applications and Innovations*, Dordrecht: Kluwer Academic Publishers, pp.243–255.

Harrer, A. (2004b) 'Soziale netzwerkanalyse in virtuellen lerngemeinschaften', in A. Martens and A. Harrer (Eds.) *Proceedings of the Workshop on Teaching and Learning Systems – The Role of Artificial Intelligence in Past, Present and Future, German Conference on Artificial Intelligence*, pp.63–72.

Harrer, A., McLaren, B.M., Walker, E., Bollen, L. and Sewall, J. (2005) 'Collaboration and cognitive tutoring: integration, empirical results and future directions', *Proceedings of the International Conference on Artificial Intelligence in Education (AIED).*

Haythornthwaite, C. (1999) 'Collaborative work networks among distributed learners', in R. Sprague (Ed.) *Proceedings of the 32nd Hawaii International Conference on System Sciences*, Los Alamitos, CA, USA: IEEE Computer Society.

Höök, K., Munro, A. and Benyon, D. (2002) *Designing Information Spaces: The Social Navigation Approach*, Berlin: Springer Verlag.

ICALTS (2004) *Kaleidoscope Network of Excellence – 'ICALTS' Interaction and Collaboration Analysis Supporting Teachers and Students Self-regulation: Deliverable 2 – Unified Framework on Interaction Analysis*, http://www.rhodes.aegean.gr/LTEE/KALEIDOSCOPE-ICALTS/Publications/D2%20Unified%20Framework%20ICALTS%20Kal%20www.pdf (accessed 7 August 2006).

Kay, J. and Kummerfeld, R.J. (1994) 'An individualised course for the C programming language', *Proceedings of the 2nd International WWW Conference*, Chicago, IL, USA: Elsevier.

Kim, A.J. (2000) *Community Building on the Web: Secret Strategies for Successful Online Communities*, Berkeley, CA, USA: Peachpit Press.

Kinshuk, L.T., Yang, A. and Patel, A. (2003) 'Plug-able intelligent tutoring and authoring: an integrated approach to problem-based learning', *International Journal of Continuing Engineering Education and Life-long Learning*, Vol. 13, Nos. 1–2, pp.95–105.

Madanmohan, T.R. and Navelkar, S. (2004) 'Roles and knowledge management in online technology communities: an ethnographic study', *International Journal of Web Based Communities*, Vol. 1, No. 1, pp.71–89.

Madrazo, L. (2003) 'SDR: networking', in B. Wasson, S. Ludvigsen and U. Hoppe (Eds.) *Proceedings of the Conference on Computer Support for Collaborative Learning*, Dordrecht: Kluwer Academic Publishers, pp.255–264.

Martínez Monés, A., Dimitriadis, Y., Rubia Avi, B., Gómez Sánchez, E. and de la Fuente Redondo, P. (2003) 'Combining qualitative evaluation and social network analysis for the study of classroom social interactions', *Computers and Education*, Vol. 41, No. 4, pp.353–368.

McLaren, B.M., Walker, E., Sewall, J., Harrer, A. and Bollen, L. (2005) 'Cognitive tutoring of collaboration: developmental and empirical steps toward realization', *Proceedings of the International Conference on Computer Supported Collaborative Learning*.

Moreno, J.L. (1951) *Sociometry, Experimental Method and the Science of Society*, New York: Beacon House, Inc.

Pinkwart, N. (2003) 'A plug-in architecture for graph based collaborative modeling systems', in U. Hoppe, F. Verdejo and J. Kay (Eds.) *Proceedings of the 11th Conference on Artificial Intelligence in Education*, Amsterdam: IOS Press, pp.535–536.

Pinkwart, N., Harrer, A., Lohmann, S. and Vetter, S. (2005) 'Integrating portal based support tools to foster learning communities in university courses', in V. Uskov (Ed.) *Proceedings of the 4th International Conference on Web-Based Education*, Anaheim, CA: ACTA Press, pp.201–206.

Preece, J., Abras, C. and Maloney-Krichmar, D. (2004) 'Designing and evaluating online communities: research speaks to emerging practice', *International Journal of Web Based Communities*, Vol. 1, No. 1, pp.2–18.

Reffay, C. and Chanier, T. (2003) 'How social network analysis can help to measure cohesion in collaborative distance learning', in B. Wasson, S. Ludvigsen and U. Hoppe (Eds.) *Proceedings of the International Conference on Computer Support for Collaborative Learning*, Dordrecht: Kluwer Academic Publishers, pp.343–352.

Sauter, W., Sauter, A. and Bender, H. (2004) *Blended Learning*, Neuwied: Luchterhand.

Scardamalia, M. (2004) 'CSILE/KnowledgeForum', in A. Kovalchick and K. Dawson (Eds.) *Education and Technology: an Encyclopedia*, Santa Barbara: ABC-CLIO, pp.183–192.

Wassermann, S. and Faust, K. (1994) *Social Network Analysis: Methods and Application*, Cambridge: University Press.

Whittaker, S., Terveen, L., Hill, W. and Cherny, L. (1998) 'The dynamics of mass interaction', *Proceedings of CSCW'98*, New York, NY: ACM Press, pp.257–264.

Xenos, M., Avouris, N., Komis, V., Stavrinoudis, D. and Margaritis, M. (2004) 'Synchronous collaboration in distance education: a case study in a computer science course', in Kinshuk *et al*. (Eds.) *Proceedings of International Conference on Advanced Learning Technologies ICALT*, Los Alamitos, CA: IEEE Computer Society, pp.500–504.

Notes

1 http://www.wiki.org

2 http://graphviz.org