Communication Patterns in Collaborative Software Engineering Courses: A Case for Computer-Supported Collaboration

Antti Knutas
Lappeenranta University of Technology
P.O. Box 20
FIN-53850 Lappeenranta, Finland
+358-0294-462-111
antti.knutas@lut.fi

Jouni Ikonen
Lappeenranta University of Technology
P.O. Box 20
FIN-53850 Lappeenranta, Finland
+358-0294-462-111
jouni.ikonen@lut.fi

Jari Porras
Lappeenranta University of Technology
P.O. Box 20
FIN-53850 Lappeenranta, Finland
+358-0294-462-111
jari.porras@lut.fi

ABSTRACT
Collaboration has become an important teaching method in software engineering and there are several computer supported collaboration tools to aid the development and learning process. However, most studies have concentrated on intra-group studies. We believe that computer supported collaborative learning tools can also aid software engineering students to have beneficial inter-group collaboration. In this research the communication patterns in three collaborative software engineering courses were analyzed with the method of social network analysis. It was found out that students do collaborate, but mostly along pre-established social connections. The main reason for this was the difficulty in matchmaking and discovering others who were struggling with the same problems. Our proposal is to study how students in similar learning scenarios benefit from computer supported collaborative tools that increase networking opportunities. The findings presented in this paper provide a baseline for comparison when performing social network analysis in future studies.

Categories and Subject Descriptors
K.3.1 [**Computers and Education**]: Computer Uses in Education – *collaborative learning.*

K.3.2 [**Computers and Education**]: Computer and Information Science Education – *computer science education.*

General Terms
Measurement, Documentation, Experimentation, Human Factors.

Keywords
Collaborative learning, computer supported collaborative learning, social network analysis, software engineering education

1. INTRODUCTION
University education and information technology teaching are going through a time of change. Learning is changing to be more interactive and the importance of collaborative learning and teamwork has grown [21]. At the same time intensive courses and team-based rapid development methods are growing more popular in software engineering education. In these approaches the goal is not only to have the students cooperate in groups, but to help each other achieve their learning goals by collaborating, for example by sharing newly learned knowledge with each other and then applying it to improve their group work. These methods have been proven to work in tertiary level education in both domestic and international studies [8, 25, 26].

However, there has been little research in how the presence of multiple groups in the same workspace affects the patterns of collaboration and if there are methods of arranging the group work in a manner that groups can benefit from each other's presence. There has been research into computer supported collaborative learning (CSCL) tools that enhance communication within classrooms, but again little research in inter-group communication. This suggests towards there being a research gap in applying CSCL learning tools to inter-group collaborative learning, which we believe to be possible and beneficial.

Collaborative learning in intensive courses (the Code Camp course series) has been studied previously in Lappeenranta University of Technology [2, 25, 26] and collaborative teaching courses in general have been studied generally [16, 24], but there have been no detailed research into inter-group communication. From observations in earlier studies it can be seen that intra-group communication occurs and some students feel that it is a beneficial part of the course [2], but more exact communication patterns and how they affect information traversal in student groups is still unclear. This gave the motivation to perform a more rigorous study on how the student communication actually occurs during the courses. Three Code Camp courses were selected for observation in order to map and analyze the student communication. The main research questions are:

1. How students utilize different communication channels during courses for collaboration?
2. Which kinds of patterns of collaboration emerge during the course, especially between different student groups?
3. Are there any available resources present in the classroom environment that could be changed to encourage more comprehensive communication or cooperation?

In order to identify emerging communication patterns in student communication in the course, the student communication that occur during courses must be mapped first. Information for the

Pre-print –version; Please cite the published paper available at
http://dl.acm.org/citation.cfm?id=2526987
doi>10.1145/2526968.2526987

study was gathered with individual surveys, recording time-lapse video for analysis and team interviews, after which patterns of collaboration were analyzed by modeling the communication patterns with the help of graph theory. The study aims to find repeating patterns with the help of social network analysis and to identify communication patterns, which could be improved with the help of CSCL tools. If any repeating patterns are found, this study can provide requirements for the next step of research, which is to implement a new CSCL system to address any found issues. The research results from this study also provide a baseline for comparison when social network analysis is applied to the improved courses.

The rest of the paper is organized as follows: the section two covers related research, the section three covers the research methods, how the research was performed and the research results, the section four analyses and discusses the results and the section five presents the conclusions.

2. COLLABORATIVE LEARNING METHODS IN SOFTWARE ENGINEERING EDUCATION

Collaboration has become an important subject of in education [13] and has been established as an essential part of 21st century skills [6]. In a collaborative learning environment, students can experience new approaches to thinking from their peers and can obtain a clearer perspective of a topic by expressing their understanding [11]. In a study by Chen et al. [7] computer supported collaborative learning tools were applied to a collaborative classroom setting, where both face to face and electronic interactions occurred, the tools used enhanced classroom communication across different groups. The intertwining of online and face-to-face collaboration was shown to unify and strengthen the collaborative learning experience [7]. However, care must be taken in applying the CSCL (computer supported collaborative learning) tools in classroom, taking into account existing social structures and pedagogical approaches [3].

Collaboration in general has been shown to be beneficial to learning software engineering [25] and that it improves both the motivation and student performance in learning programming [19]. The use of computer supported collaborative learning tools specifically in software engineering education courses is also established and it has been shown that students can collaborate directly using computer collaboration tools [9] with beneficial results, share information using messaging and annotation services [17] and collaborate in groups using groupware software [5] or wikis [20]. However, we found no evidence of studies concentrating on the research of how introducing CSCL tools affect inter-group interactions. This research gap could be addressed by researching how collaboration social networks in the classroom function and if there are any problem spots that could be addressed with the introduction of CSCL tools.

The problem of analyzing classroom interaction can be approached with social network analysis. Social network analysis (SNA) is an interdisciplinary technique for the analysis of social networks [22], where social relationships are viewed in the terms of network theory. In social network analysis communication between individual or social units are mapped into a communication matrix and then visualized in graphs. In graph theory there are different mathematical tools available, which can be used to for example estimate the relative influence of nodes in the graph or analyze the graph by the nodes' connection patterns [1, 14]. In this research case the communication patterns of different groups could be analyzed by modeling students as nodes and mapping student and intra-group communication as node edges.

This kind of analysis has been applied to collaborative learning by [27] to model collaboration in distance learning groups and was applied with additional qualitative analysis to CSCL classroom learning scenarios [18]. They found out that it is possible to apply SNA to CSCL scenarios, but Martinetz et al. [18] concluded main difficulties were related to the speed of processing. More specifically, it is difficult to gain results fast enough to provide corrective feedback during the progress of the course. Additionally, the process depends on the expertise of the researcher and it is difficult to automate.

3. COMMUNICATION PATTERNS IN INTRA-GROUP STUDENT COLLABORATION

The study about observing student communication patterns in software engineering course was carried out during three five days long collaborative courses. At the start of the course the students were first divided into work groups and were assigned a programming task with a deadline, but were allowed rather freely to decide how to achieve their goal. While the students worked towards their goal, their inter-group communication was observed both by a researcher and by automated tools, which was later analyzed with SNA analysis methods.

The Code Camp style intensive courses last five days, with the course structure clarified in the Figure 1. The course starts with a technical introduction and task assignment in a traditional lecturing environment. Usually at this point no groups are formed, though there is some informal chatter between some students and some study groups have been known to agree to form a group. After the first day's lectures and the task assignment, a social event is held where the students are allowed to get to know each other more and form groups for the next day. The following days concentrate on actual implementation work and during the last day there are student presentations and a friendly competition between the project works.

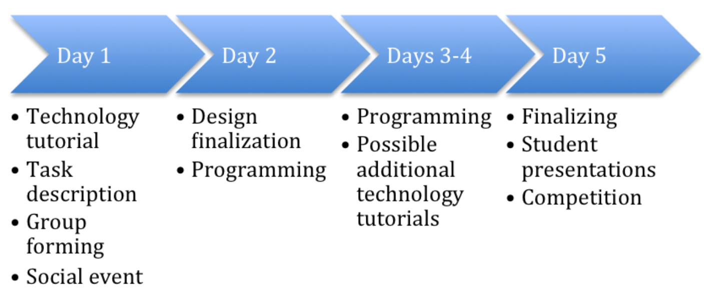

Figure 1. Code Camp course outline

During the course most of the information was first made available through a lecture, but the course had a supporting wiki page for distributing and coordinating information. Course material was made available there, including lecture slides, software tools and schedules. All students had edit access for the wiki and there was a course requirement of each team creating a wiki page for group coordination. From the student group's point of view these wiki pages were used as communication hubs, with team members updating their project status and uploading course deliverables there. The wiki had no access restrictions; so all participants of the course were free to examine all available materials.

There were two levels of support during the course: Collaborative and teacher-provided. During the course, the students and the groups were allowed to freely communicate with each other and share code to solutions with face-to-face communication or with online tools. In addition to being present in the classroom constantly during the course, the teachers entered most often asked questions into the Q&A wiki page, linked to code examples and uploaded course materials like lecture slides.

3.1 Research Setup

The observation of the communication patterns was performed by three different methods: Direct observation of the students by having one researcher present in the classroom and observing the conversations, time lapse video monitoring of the groups interacting in the classroom and recording wiki activity on the web server. After the course had ended, both the students and the teaching personnel were asked to fill a survey about their usage of resources to gain additional data to supplement the direct observations. Additionally in the second and third observed courses the teams were interviewed using a qualitative approach in order to gain better insight to the interactions and to try to detect interactions that were considered but were not taken. The questions that were asked in the interviews were:

- What went well?
- What didn't go well?
- How did you cooperate with other teams?
- Were there any reasons why you were not able to collaborate with the other teams?

The interviews were performed using a semi-structured format. When the teams answered to the questions, further questions based on the teams' answers were presented, especially if they mentioned any factor that affected teamwork or communication positively or negatively. After the graph visualization was completed, the interviews were studied in order to gain perspective how each individual group perceived the effect of communication on their work. Additionally, any factors that affected communication, intergroup cooperation or teamwork were noted and collated.

The main source for the interaction data was video monitoring. The classrooms were monitored with two web cameras that covered the observed groups and captured single images at a pace of an image per second, or six images per second when motion detected in the room. When these single images are converted into video frames they form a time-lapse video that allows the entire five day event to be viewed at a rapid speed. The following data was recorded from the video for further analysis:

- Time of communication
- Type of event (lecture, active group work, break)
- Initiating group member
- Receiving group (or teacher)

The following additional data was gathered with the surveys:

- Group the student most often preferred to collaborate with
- How often the student asked the assistant for help
- How often the student used online resources
- How often the student collaborated with other teams

After each course the videos were first analyzed and the interactions recorded and then collated. The surveys and researcher's observation notes were used to set context for each interaction. After collating the interactions graph theory and social network graphing software are used to form social network interaction graphs of the physical interactions occurring in the classroom.

3.2 Methods of Analysis

The patterns of collaboration were analyzed by modeling the communication patterns with the help of graph theory. Each interaction, the interaction context and reason for the interaction were recorded from the available raw material. The lists of interactions were collated into a directed graph, where the nodes represent individual students and the edges represent communications between the nodes. The graph was analyzed by inputting it into the graph analysis software Gephi [4] and using the visualizations produced by the software to identify influential nodes, strength of cooperation between groups and repeating patterns of collaboration between the nodes. The analysis software uses the Force-Atlas algorithm to map the nodes, taking into consideration the relative importance (size) and connection strengths of the nodes [10].

In order to gain more understanding into the arrangements of the nodes into the graph, influence analysis was performed on the graphs using the PageRank indexing algorithm [23], which can also be used to measure influence of nodes in social networks [12, 28]. For example the influence of Twitter users has been researched with this SNA approach [15, 29]. The algorithm calculates a probability distribution for arriving to a specific page in a graph representing the links of a set of hypertext documents. The values returned by the algorithm are normalized so that the sum of values is 1. An essential part of the PageRank algorithm is the dampening factor, which causes pages that have few links to important pages to be valued more highly than pages that have a wider array of more random links [23].

The PageRank algorithm is expressed in the Equation 1, where the $PR(A)$ is the PageRank of Page A, $PR(Ti)$ is the PageRank of pages Ti which link to page A, $C(Ti)$ is the number of outbound links on page Ti and d is a dampening factor which can be set between 0 and 1, but is usually set to 0.85. PageRank or $PR(A)$ can be calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the web.

$$PR(A) = (1-d) + d\,(PR(T1)/C(T1) + \ldots + PR(Tn)/C(Tn))$$

Equation 1. The PageRank algorithm. [23]

3.3 Observed Communication Patterns

Three courses were observed during a one-year period, with a total of 42 students spread across fifteen student groups. The studied courses were optional and attracted students with varying backgrounds from different majors, but are generally master's level students.

In the first course eight groups were formed at the start of the course, seven of which completed the course and were selected for analysis for a total of twenty students. 44% of the participating students had attended an Code Camp –style events before this course. One group was left out from the study, because the members left the course before the end of the second day and did not participate in the activities.

Students actively collaborated both inside their groups and between the groups during the observed courses. The groups formed communication patterns that were modeled into graphs and are presented later in this section. Each node represents a student and the edges connecting the nodes represent collaboration-related communications that occur between each individual student. The thicker the connecting line in the figure, the more often the group member participated in the communication, with the relative strength of the communication rated from one to three. For example, in the Figure 2 the connection between D1 and D3 is level three, D3 and A1 connection level two and the D1 – F1 connection level one. The relative sizes of the nodes are based on the PageRank algorithm, which is a measure of a node's influence in a graph. The colors and the sizes of the nodes, from blue to red and small to large, indicate the nodes' relative PageRank values.

All of the seven student groups who completed the course were analyzed further. The communication patterns for the groups in the first course are presented in the Figure 2. An alphabet depicts each group and a number each group member. For example D3 means member 3 in the group D. The different student groups formed communication clusters, with certain groups having stronger communication ties with each other. It can be seen that the nodes D1, D3, A3 and C3 forms an especially important communication center, relaying often information between the nodes' own group and two other groups. This correlates with the observer's observation that the D group was influential and several other members visited the table, asking for programming advice. The student D3 also visited his friends' tables, relaying information about recent solutions. The other noticeable pattern is that one student is particularly active in communicating with other groups and relaying the information to their own group, as in the case of nodes F2 and B1. It can also be seen that some groups overall form communication hubs, becoming central to the graph. The groups A and D are influential in the classroom communication social network, with the E and C group acting as intermediaries for information.

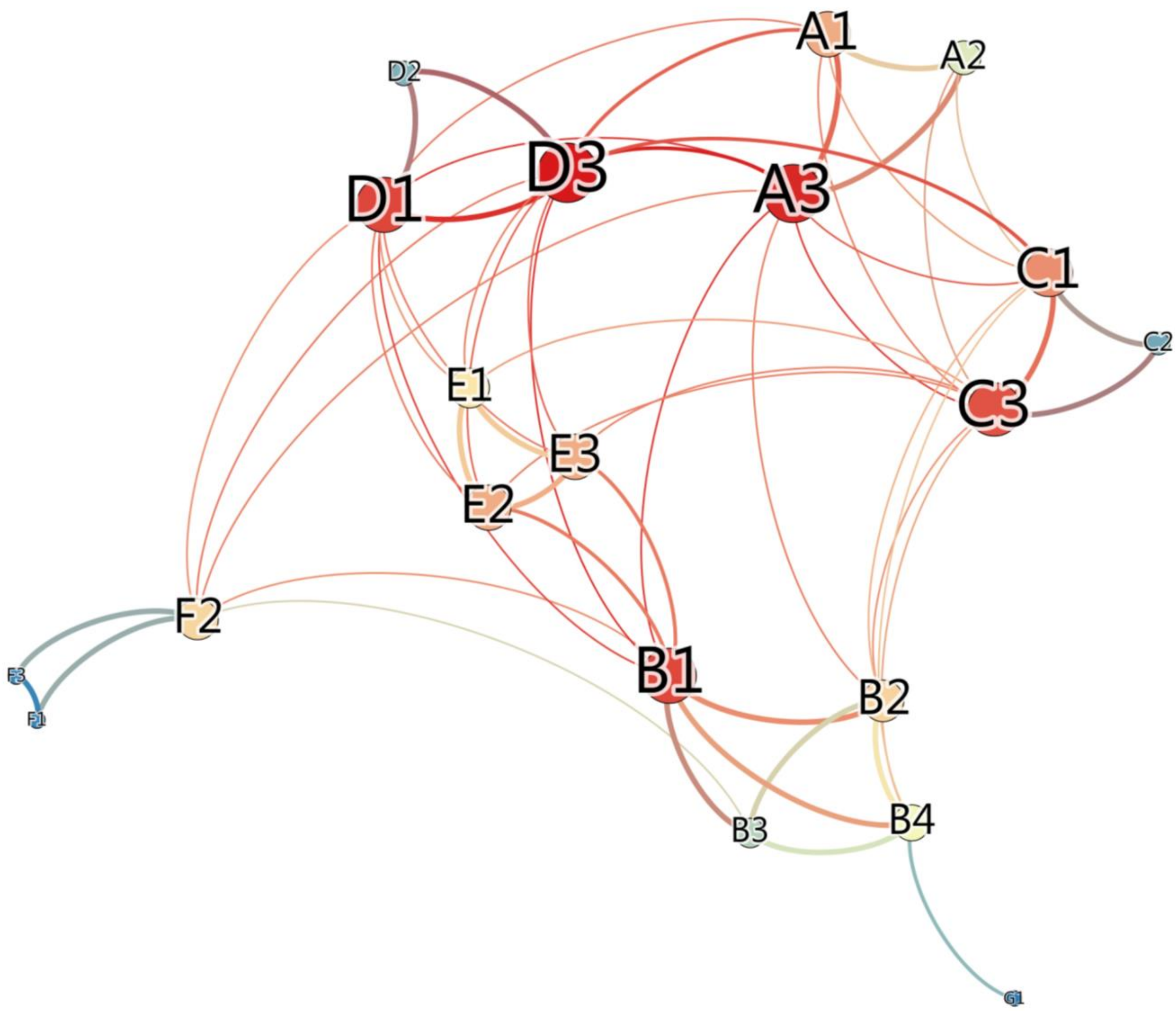

Figure 2. Student group communication graph, first course

The participants were polled for use of other communication resources, in addition to observing their classroom communication patterns. The communication frequencies are presented in the Table 1. The communication frequencies are collated to three levels: Rarely, daily or several times per day. Similarly the course success is divided to three tiers: Top third, the middle tier and last third. It can be seen that some of the most influential groups communicated a lot with the assistant and from observing the classroom it could be seen that often the information given by the assistant was spread to other students through physical communications that followed the presented communication network. All groups used online resources often in addition to communicating with each other.

Table 1. Group communication frequencies, first course

Group	Asked for help from the assistants	Used online materials	Communicated for ideas or technical from other groups	Course success tier
A	Several times per day	Several times per day	Several times per day	2nd best
B	Daily	Several times per day	Daily	2nd best
C	Rarely, once or twice	Several times per day	Several times per day	2nd best
D	Several times per day	Several times per day	Several times per day	1st
E	Rarely, once or twice	Several times per day	Several times per day	2nd best
F	Rarely, once or twice	Daily	Daily	2nd best
G	Several times per day	Several times per day	Several times per day	3rd best

The second course had five student groups, four of which were selected for observation for a total of sixteen students. 24% of the participating students had attended a Code Camp –style event before this course. One student group was excluded because they only partly attended the event. In the second course the communication graph, presented in Figure 3, forms a circle instead of a set of clusters. Two major groups, the groups A and C, form a tight and influential cluster, which mostly connect to two separate groups. Again the tight cluster correlates with tight sociability: In the social event the groups got to know each other better and the students A3, A2, C2 and C4 spent a good part of discussing during the social event. However, when inspected from the perspective of utility, the collaborations are not as effective as they could be: The groups C and D both worked on the same mobile platform and collaborated little, while the groups A and C collaborated on completely different concepts and platform. Similarly, the groups B and C worked on similar problems and only some of the students collaborated. As in the previous course, some groups, A and C in this case, form communication cluster and in some groups there is a distinct person, like D1 or B4, who communicates with this influential cluster.

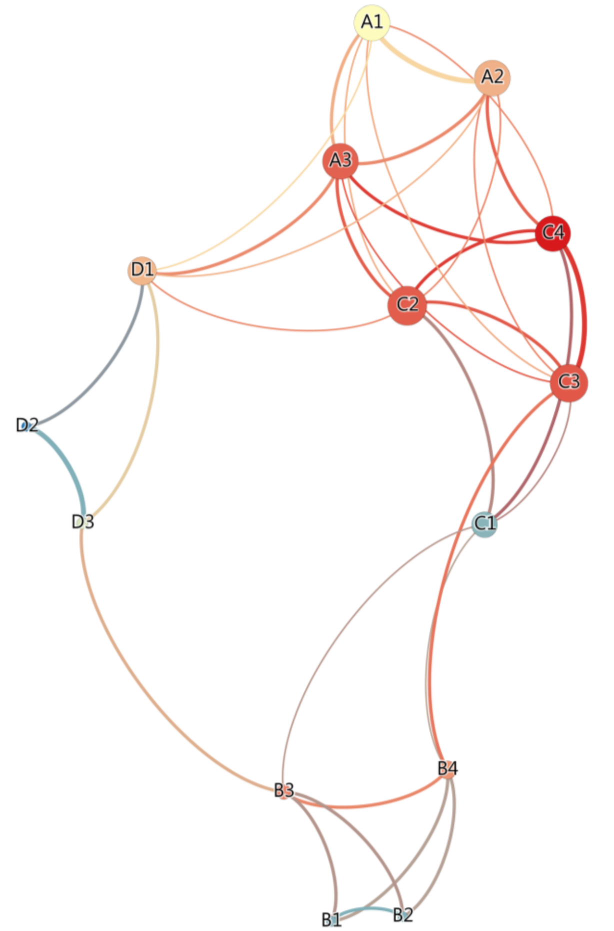

Figure 3. Student group communication graph, second course

The students participating for the second course were also polled. The communication frequencies for the second course are presented in the Table 2. The course success was very similar for all participating groups in this instance, with difficulties in finding clear differences between groups. Unlike in a previous course, the communication between groups was less frequent and one influential group reported that they received ideas and technical help from other groups less often. The other influential group C repeated the pattern observed in the first course and communicated often with the course assistant, sharing technical help with the other groups. Of all the groups, the groups B and C were physically closest to each other.

Table 2. Group communication frequencies, second course

Group	Asked for help from the assistants	Used online materials	Communicated for ideas or technical from other groups	Course success tier
A	Daily	Several times per day	Rarely, once or twice	2nd best

B	Rarely, once or twice	Several times per day	Daily	3rd best
C	Daily	Several times per day	Daily	3rd best
D	Rarely, once or twice	Several times per day	Rarely	1st

In addition to asking for numerical answers, the poll was expanded with optional textual answers asking about most important communication factors that helped the entire group to progress. The answers included good communication amongst the team, resourceful team members, meetings, online help and resources, friendly working environment, advice and ideas from other groups and sharing information. The most commonly mentioned sentences were related to information sharing and friendly working environment or team spirit.

The student groups participating in the second course were interviewed as teams in addition to being polled. The most comments involving collaboration mentioned it being allowed to collaborate with other groups useful, but difficult to initiate. Several students mentioned that collaboration was difficult to initiate with more distant groups, because it was difficult to know what they were currently working on. They said that another issue was about knowing when to establish collaboration. Programming is mostly a quiet activity and it was difficult to tell when the person did not want to be disturbed. However, at the same time several interviewees said that they could have welcomed more requests for collaboration, but other people did not initiate them.

The third and the final studied course had fewer students, because of its position at the end of the semester and had only three student groups, all of which were selected for observation for a total of seven students. 80% of the students in this course had participated to Code Camp –style events before, with overlap from the second observed course. The communication graph that student collaboration communication forms is presented in the Figure 4. Again in here it can be seen that the most influential students form a tighter communication cluster, with one student and one group at the edges. However, some students are more left out from the tight core of communication, with only faint and occasional instances of collaboration. It should be noted that the students A2, B1 and B2 correspond to students A1, A2 and A3 in the previous code camps. It is a pattern that can also be seen repeating in other courses that are outside this study: Students tend to network and tend to collaborate with each other, whether formally in the same group or not.

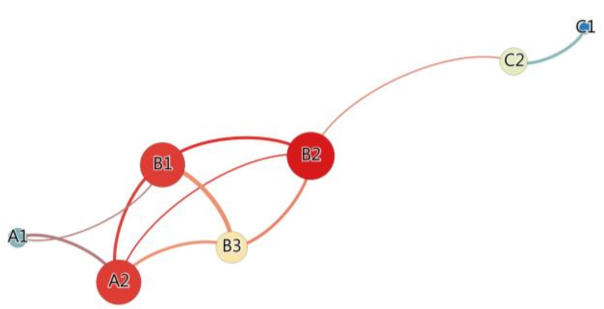

Figure 4. Student group communication graph, third course

The communication frequencies for the third course are presented in the Table 3. While the course had fewer participants than the previous ones, the differences in success were more noticeable. It can be seen from the observations that the most successful groups, A and B, used online resources often in addition to communicating with each other.

Table 3. Group communication frequencies, third course

Group	Asked for help from the assistants	Used online materials	Communicated for ideas or technical from other groups	Course success tier
A	Rarely, once or twice	Several times per day	Daily	2nd best
B	Daily	Several times per day	Daily	1st
C	Daily	Daily	Rarely, once or twice	3rd best

The textual part of the poll asked about most important communication factors that helped the entire group to progress. Student replies included team spirit, other people, shared ambition, working cooperation, version control software and online resources. Cooperation in work and team spirit were mentioned most often.

Like in the second course, students participating in the third course were team interviewed. The interview results were similar as well: The major difficulties to communication according to team interviews were the diversity of topics and the difficulty in establishing contact to people the students did not know beforehand.

Finally, the usage of communication resources was observed and polled in all three courses in aggregate. As presented in the Table 4, the overall usage of all communication resources in all of the courses was high. Most of the resources were in usage daily or more often, except for group-to-group communication in the second course. A possible explanation for this was the diversity of allowed topics, so that the groups had less to contribute to each other.

Table 4. Aggregate communication frequencies

Course	Asked for help from the assistant daily or more often	Used online materials daily or more often	Communicated with other groups daily or more often
1st	61%	100%	67%
2nd	50%	100%	33%
3rd	71%	100%	71%

3.4 Analyzing the Communication Patterns

From the different communication patterns it can be seen that the students collaborated a good amount, both within the groups and to a certain extent between the groups. There were repeating patterns discovered in the student communications. First observation is that students willingly collaborate between groups despite a wealth of available online resources and a slightly competitive atmosphere being present in the classroom. When interviewed and observed in the classroom and the social events, it was revealed that often the first or strongest collaboration occurs between people who know each other from previous social contexts. This collaboration can expand when more people see this center of activity and join in. The second most common provided reason for initiating communications was pre-established knowledge that the person might have useful information or be able to help with a problem.

The use of online resources, collaboration networks and communication frequencies between the groups differed between the three courses. The nodes in the first collaboration graph formed loosely connected clusters, while in the second course the pattern was a rough circle, with less interconnectedness but stronger collaboration in a single cluster. The probable reason for this is the difference in the task assignment: In the first course all the groups were assigned a similar task on a same platform and in the second course the groups were allowed to pick their own implementation, platform and topic within a specific theme. The same reason can explain higher communication frequency between groups in the first course. Lastly, in the third course there was a tight cluster of collaboration that was as strong as the collaboration inside of one of the groups. In interviews the provided reason for this collaboration was existing friendships between many of the collaborators from different groups. It appears that a certain minimum amount of groups is required in order to have varied collaboration between groups. One possible pattern that did not occur in the study is one in which two strong competing centers of cooperation form, with the groups essentially split in two separate blocs.

In all of these courses one or two core groups of collaborators were discovered. In the first course they were groups A and D, in the second course groups A and C and in the third course groups A and B. These groups are characterized by a certain enthusiasm, willingness to communicate and were well graded and placed well at the competition that was held at the end of the courses. The groups that communicated often with each other also used other communication resources, like online materials both on the course pages and other web materials. Other groups benefited from interactions with them, since the groups' willingness to communicate extended to sharing information they had discovered. However, not all groups were equally connected to these strong centers of collaboration and did not equally benefit from it.

4. DISCUSSION ON CLASSROOM COLLABORATION COMMUNICATION PATTERNS

Our study showed that overall the communication patterns appear to follow pre-existing, physical social networks that were established outside the context of the course. The question is whether this is an optimal case and whether the classroom or online environments can be changed to encourage forming new, beneficial social connections. One of the major issues that were mentioned in the team interviews was the difficulty of knowing whether it is appropriate or beneficial to establish communications. This is the major reason that inter-group collaboration follows pre-existing social connections, because the barrier to establish communication is less difficult. Informal communication channels between friends also provide more information about group project work status, which makes it easier to initiate collaboration.

The analysis was performed using three different approaches, which are video and direct observation, interviews and polls. Results from the different sources were used to complement each other. For example, direct observation and interviews provided context data while the video observation covered the event completely from start to finish for most accurate graphing data. Student participation in the data gathering was high, with every observed group participating in the interviews in the last two courses and 95% of the students completing the online surveys. With the high percentage of data collection and accurate observation records from the entire duration of the events it can be said that the structure of the graphs are an accurate abstraction of what occurs in the classroom. The study would have been more conclusive if it had included more courses in a wider range of universities. However, the studied communication channels in the three observed courses already provided notable patterns of communications for analysis. Also, some of the patterns repeated during each of the courses, suggesting towards them being common patterns to Code Camp –style courses in environments that are similar to the study environment.

The main pattern that repeated in all of the courses was a strong center of collaboration that formed around one to three groups. The fact of strong center of collaboration was contrasted with the interview reports of difficulties in matchmaking when looking outside one's group. This suggests that not all groups find or are not able to work with the strong center of collaboration and that collaboration does not equally benefit all participating groups.

We propose that the discovered issues can be addressed by offering computer supported collaborative tools that support wide student-initiated collaboration. A common issue during the courses was that the students did not realize that they had similar problems, which caused hesitation in initiating communication. Software tools can be used to publicize commonly encountered problems and to find people who are struggling with the same problems for collaboration. Tools like these could help students find each other without spending time on discovering partners and accidentally disturbing people who are concentrating on individual problem solving. An additional benefit would be that the problem and the following conversation would be recorded for other participants to view later in the course if they struggle with a similar problem. For example question and answer sites with reward systems have seen wide use in the field and could be also applied inside classroom. Additional tools, like projectors or mobile clients, could be used to publish unanswered questions and the most useful solutions.

Traditional online courseware tools are now commonly seeing use in classroom environments, but their usage focus is often to provide course literature, assignments and accept returns. While they do allow things like peer review of assignments, this style of collaboration is teacher controlled and usually more slowly paced. This study shows that the patterns of collaboration in the Code Camp style of courses could be improved and that additional tools should be provided to encourage student-initiated collaboration. Improved software tools could give more opportunities for collaboration within classrooms by easing sharing, providing matchmaking services and should be investigated in future research. There are several examples of web-based collaborative tools widely used in the industry. One is Github, which enables direct collaboration between several people by allowing several people to edit and version software source code concurrently. Another is Stack Overflow, which facilitates problem solving by allowing people to post questions and rate and reward the best answers with reputation points.

Computer-based communication tools do require the presence of computers in the learning environments and this could a drawback in adopting the tools. However, in software engineering courses computers are already present as development tools and using computer-based collaboration tools will most likely have a lower barrier for adoption than in other fields of education. Also, software engineering students are already familiar with using online tools for working. There is a possibility that their current development or planning tools do not have to be replaced, just extended. There are plugins or workflows that support collaboration and collaborative learning both within and between groups and these tools can be introduced into the course workflow and be pre-installed into the working environments.

5. CONCLUSION

In this study we applied social network analysis to intensive collaborative software engineering courses using recordings, polls and interviews as source material. We presented and analyzed the communication collaboration patterns that form during intensive collaborative software engineering courses. It was found out that the students do collaborate outside their groups on problems, but the patterns of collaboration follow pre-established social connections and not all groups equally benefit from the collaboration. The main method of collaboration was seeking out these social connections, like well-known classmates or friends and discussing with them, whether they were working on the same problem or not. The main result of this study is discovering the form of communication patterns that are established during the courses. These patterns and discovered issues in matchmaking can provide the basis for designing CSCL tools to improve collaboration. Additionally the results can be used to validate and compare improvements to communication patterns when applying social network analysis to future courses that use CSCL tools.

While students mostly collaborate along pre-existing social connections, almost all of the groups in the observed courses used a major amount of online resources and used computers for planning from the start. Our proposed solution for improving collaboration in these already computer-supported work processes is introducing of online collaboration tools and groupware solutions to that are already well established elsewhere in the industry, instead of the more often used classroom online tools. These online collaborative tools, such as Stack Overflow or Github, can fit the fluid nature of the event better. The classroom tools are often aimed for delivering preplanned course material according to a curriculum.

During the courses students mainly communicated face to face, but because of privacy and monitoring issues there was no possibility to gain detailed information about the students' electronic communications. However, in further social network analysis research where a centralized CSCL tool is designed and used for comparison, these limitations can be overcome by collecting statistics from the collaborative tools.

6. REFERENCES

[1] Abraham, A. and Hassanien, A.E. *Computational Social Network Analysis: Trends, Tools and Research Advances*. Springer, 2010.

[2] Alaoutinen, S. et al. Experiences of learning styles in an intensive collaborative course. *International Journal of Technology and Design Education*. 22, 1 (2012), 25–49.

[3] Baker, M. et al. Integrating computer-supported collaborative learning into the classroom: the anatomy of a failure. *Journal of Computer Assisted Learning*. 28, 2 (2012), 161–176.

[4] Bastian, M. et al. Gephi: An open source software for exploring and manipulating networks. *International AAAI conference on weblogs and social media* (2009).

[5] Bravo, C. et al. Integrating educational tools for collaborative Computer Programming learning. *Journal of Universal Computer Science*. 11, 9 (2005), 1505–1517.

[6] Bruns, A. Produsage. *Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition* (New York, NY, USA, 2007), 99–106.

[7] Chen, W. et al. What do students do in a F2F CSCL classroom? The optimization of multiple communications modes. *Computers & Education*. 55, 3 (Nov. 2010), 1159–1170.

[8] Davies, W.M. Intensive teaching formats: A review. *Issues in Educational Research*. 16, 1 (2006), 1–20.

[9] Duque, R. and Bravo, C. Analyzing Work Productivity and Program Quality in Collaborative Programming. *The Third International Conference on Software Engineering Advances, 2008. ICSEA '08* (2008), 270–276.

[10] Force-Atlas Graph Layout Algorithm: URL *http://www.medialab.sciences-po.fr/publications/Jacomy_Heymann_Venturini-Force_Atlas2.pdf* (2012). Accessed 12 June 2013.

[11] Gillies, R.M. Teachers' and students' verbal behaviours during cooperative and small-group learning. *British Journal of Educational Psychology*. 76, 2 (2006), 271–287.

[12] Java, A. et al. Modeling the spread of influence on the blogosphere. *Proceedings of the 15th international world wide web conference* (2006), 22–26.

[13] Johnson, D.W. and Johnson, R.T. Learning Together and Alone: Overview and Meta- analysis. *Asia Pacific Journal of Education*. 22, 1 (2002), 95–105.

[14] Knoke, D. et al. *Social network analysis*. Sage Publications, second edition, 2007.

[15] Kwak, H. et al. What is Twitter, a social network or a news media? *Proceedings of the 19th international conference on World wide web* (New York, NY, USA, 2010), 591–600.

[16] Laat, M. de et al. Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for Social Network Analysis. *International Journal of Computer-Supported Collaborative Learning*. 2, 1 (Mar. 2007), 87–103.

[17] Lan, Y.-F. and Jiang, Y.-C. Using Instant Messaging and Annotation Services to Improve Undergraduate Programming Courses in Web-Based Collaborative Learning. *Fifth International Joint Conference on INC, IMS and IDC, 2009. NCM '09* (2009), 236–241.

[18] Martínez, A. et al. Studying participation networks in collaboration using mixed methods. *International Journal of Computer-Supported Collaborative Learning*. 1, 3 (Sep. 2006), 383–408.

[19] McDowell, C. et al. The effects of pair-programming on performance in an introductory programming course. *SIGCSE Bull.* 34, 1 (Feb. 2002), 38–42.

[20] Minocha, S. and Thomas, P.G. Collaborative Learning in a Wiki Environment: Experiences from a software engineering course. *New Review of Hypermedia and Multimedia*. 13, 2 (Dec. 2007), 187–209.

[21] Okamoto, T. Collaborative technology and new e-pedagogy. *IEEE International Conference on Advanced Learning Technologies, 2004. Proceedings* (Sep. 2004), 1046 – 1047.

[22] Otte, E. and Rousseau, R. Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*. 28, 6 (Dec. 2002), 441–453.

[23] Page, L. et al. The PageRank citation ranking: bringing order to the web. 1999.

[24] Pais Marden, M. and Herrington, J. Supporting interaction and collaboration in the language classroom through computer mediated communication. *World Conference on Educational Multimedia, Hypermedia and Telecommunications* (2011), 1161-1168.

[25] Porras, J. et al. Better programming skills through Code Camp approach. *16th EAEEIE Annual Conference on Innovation in Education for Electrical and Information Engineering, Lappeenranta* (2005), 6–8.

[26] Porras, J. et al. Code camp: a setting for collaborative learning of programming. *Adv. Technol. Learn.* 4, 1 (Jan. 2007), 43–52.

[27] Reffay, C. and Chanier, T. Social Network Analysis used for modelling collaboration in distance learning groups. *Intelligent Tutoring Systems* (2002), 31–40.

[28] Scott, J. *Social network analysis*. SAGE Publications, third edition, 2012.

[29] Weng, J. et al. TwitterRank: finding topic-sensitive influential twitterers. *Proceedings of the third ACM international conference on Web search and data mining* (New York, NY, USA, 2010), 261–270.