

CORRELATING PERFORMANCE WITH SOCIAL NETWORK STRUCTURE THROUGH TEACHING SOCIAL NETWORK ANALYSIS

Peter Gloor

MIT, pgloor@mit.edu

Maria Paasivaara

Helsinki University of Technology, mpaasi@hut.fi

Detlef Schoder

University of Cologne, schoder@wim.uni-koeln.de

Paul Willems

University of Cologne, Paul_Willems@gmx.de

Teaching a course on optimizing online communication behavior and social network analysis permitted us to obtain preliminary results on correlating temporal online communication patterns with team performance. Students from Helsinki University of Technology and University of Cologne who had never met face to face formed virtual interdisciplinary teams collaborating on a common task. While collaborating over long distance, students kept track of their own communication activities by e-mail, chat, and conference calls with Skype. The contribution of this paper is twofold. First, we introduce an innovative course format creating an empirical base for team performance in a distributed online communication environment. Secondly, we provide basic analysis of correlations between SNA measures and team performance. Students used these insights to optimize their own communication behavior for future virtual collaboration.

1. INTRODUCTION

The advent of the Internet has provided new opportunities for collaboration thought impossible just a few years ago. Exchanging ideas and work by e-mail, chat, Internet telephony, blogs, and Wikis has opened up new avenues for spontaneous communication. Researchers have begun to study how these new communication channels influence productivity and creativity of virtual teams [Cro04, Cum03, Glo03, Kid05, Lee03, Lue03]. In our own work we studied Collaborative Innovation Networks, or COINs [Glo06]. COINs are virtual teams of self-motivated people with a collective vision, enabled by technology to collaborate in achieving a common goal – an innovation – by sharing ideas, information, and work.

This paper describes early results on how to improve online communication for better performance and creativity. These insights have been gained when teaching a graduate-level distributed course on online collaboration co-located at three universities. The main objective of this course had been to offer students an opportunity to improve their own communication behavior when collaborating in virtual teams to become better members of COINs. They did this by completing an innovation-centered distributed project as a virtual team, and then correlating their individual and team communication pattern with individual and team performance achieved in the distributed project.

In fall 2006 we jointly taught a course to 13 students in Finland and 12 students in Germany on how to optimize their online communication behavior to

become better net citizens and members of virtual teams, increasing their efficiency and creativity. Part of the course was taught from MIT, such that the course was distributed at three locations. Figure 1 illustrates the classroom teaching part of the course, where one virtual classroom was formed by participants from Helsinki, Cologne, and Boston.



Figure 1 – Snapshots from teaching the course

Our course was organized in three parts. In part one, students learned about principles of social network analysis (SNA) [Was94], Collaborative Innovation Networks (COINs), and Swarm Creativity [Glo06]. In the second part students formed seven interdisciplinary teams comprising three to four students from different institutions (University of Cologne, Helsinki University of Technology) and applied the tools framework taught in part one by analyzing a virtual online community. This permitted them to study rules of optimized online communication in other online communities. In part three the students analyzed the communication behavior of their own virtual team, based on their online communication record of e-mail, chat, and phone interaction. Communication records were collected by cc'ing all communication activities to seven dummy e-mail addresses.

The main goals of the course were to teach students how to be efficient online-communicators and collaborators in distributed virtual teams. Our objective was for students to become more effective communicators by becoming aware of their social position and their contribution pattern in the virtual team. In more general terms, course participants also learned to increase organizational innovation and effectiveness by converting organizations into “Collaborative Innovation Networks” (COINs). On a technical level, they learned how to apply social network analysis using the tool TeCFlow (Temporal Communication Flow Optimizer) developed at MIT and Dartmouth [Glo04].

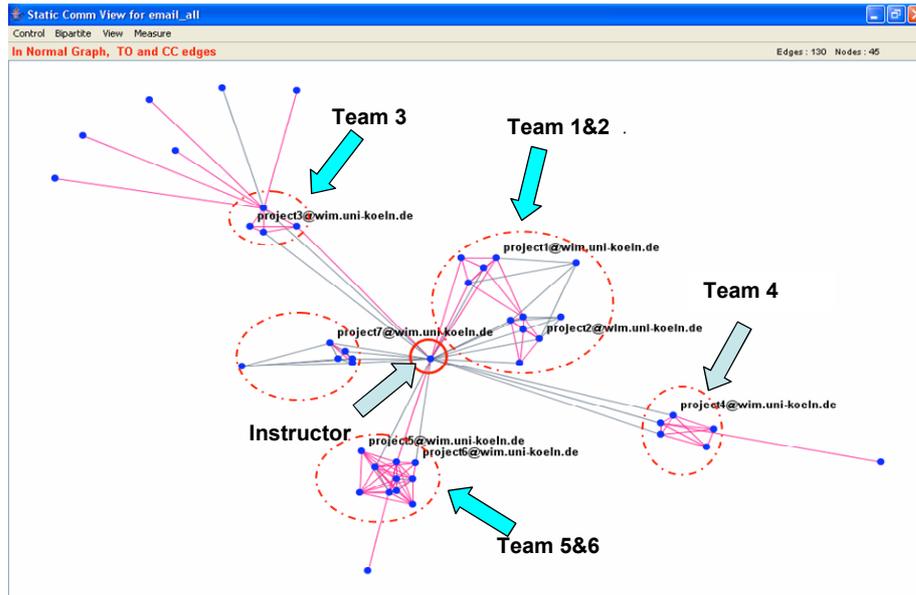


Figure 2 – Communication patterns of project teams in phase 2.

Figure 2 shows the social network during part two of the course, produced with TeCFlow. Ties between actors have been produced by mining the e-mail archive of the course communication. Note the central role of the instructor, with very little inter-team communication. Only teams 1 and 2, and teams 5 and 6 show inter-team communication.

2. CORRELATING PERFORMANCE WITH SOCIAL NETWORK STRUCTURE

The course participants formed seven separate project teams, each with team members from both Cologne and Helsinki. Each team analyzed an online community. They choose subjects such as communication among contributors to Wikinews, tracking of trends on RFID through using the ISIWeb literature database or analyzing e-mail communication among Enron employees. As the team members were geographically distributed, their communication was conducted online, mostly by e-mail.

Each student graded the quality of the work of the teams other than her or his own team on a scale from 1 to 4, with 1 being the best grade. The quality of the work of each team was ranked based on the quality of the final presentation of the team and the final report. Students also ranked the quality of the individual contribution of their own team members. This means that each student gave a grade to each of the other six teams, and to the two to three peers within the team. The best students and teams were rated 1, the worst a team was rated by a student was 3, the worst an individual was rated was a 4.

We tested three hypotheses based on the peer ratings. The three hypotheses are:

1. The internal team ratings are correlated to the communication balance of the teams.
2. The external team ratings are correlated to the communication balance of the teams.

3. There is a significant correlation between the external ranking of each team's output and the mutual internal ranking among team members.¹

We also looked at more simple parameters such as the number of e-mails sent within each team. While there was indeed correlation between external rating and numbers of messages exchanged, it turned out not to be significant. This may be because of the small size of our sample. Applying typical SNA measures such as betweenness and degree centrality [Was94] did not make sense here, because of the small individual team size of three to four members, which were all fully connected.

The hypotheses were tested on the communication data collected from the course and the grades. All e-mail communication between the course participants was collected and was used as the basis for the communication analysis. The main measure to be used for this analysis was the contribution index, which is defined as:

$$\text{contribution_index} = \frac{\text{messages_sent} - \text{messages_received}}{\text{messages_sent} + \text{messages_received}} \quad [\text{Glo03}].$$

The contribution index takes on values between -1 and $+1$, it is $+1$ if a person only sends messages and -1 if it only receives messages. A contribution index of 0 indicates a totally balanced communication behavior. In the project most project related communication was communication within the team. Peripheral communication of team members with outside people, which would have had an impact on their contribution index and might have distorted its relation to the contribution index of the other team members, was almost non-existent.

Average weighted variance of the contribution index (awvci)

$$\text{awvci}^{ws}(\text{team}) = \frac{\sum_{i=1}^T m_i \cdot \text{var}(C_i^{ws})}{\sum_{i=1}^T m_i}$$

m_i = number of edges on day i
 ws = window size
 $C_i^{ws} = (c_{i,n}^{ws}, \dots, c_{i,N}^{ws})$ = vector of contribution indexes of all team members for day i
 $c_{i,n}^{ws}$ = contribution index of member n on day i

Figure 3 – Average weighted variance of contribution index to calculate the balanced'ness of a team's communication

To describe the balance of a team's communication the average weighted variance of the contribution index (figure 3) was defined. The window size denotes the sliding time window in number of days used to calculate the contribution index with TeCFlow. In order to reduce the impact of high variances of the ci caused by single messages by one member in phases of general low activity, which leads to distorting (weekend) peaks, the variances are weighted with the number of total edges on that particular day. This weighting increases the influence of patterns that appear in high activity phases like shortly before the deadlines. The resulting average weighted variance of the contribution index ($awvci$) adopts values close to 0 if the communication is balanced.

¹Tested by Lutz Tegethoff, Ilkka Lyytinen and Sebastian Schiefer during part three of the COIN course

2.1 Teams with high internal ratings have a ‘balanced’ internal communication behavior

The average internal ratings can be seen as a self assessment of a team. Teams in which information is shared fast and even among the members and where the response time on mails is short, should have more satisfied members, which give better ratings to the rest of the team. These teams are also expected to have a low *awvci*. Therefore the average internal rating should be correlated to the *awvci*.

As it turns out, there was no significant correlation between balanced internal communication behavior and internal rating² (see Table 1). We can speculate that team members differed in their willingness to give each other “harsh” grades, thus distorting the measurements in our small sample.

n=7 teams	<i>awvci</i>			
	Window Size 1		Window Size 5	
	Pearson Correlation	P-Value	Pearson Correlation	P-Value
External Rating	0.724	0.066	0.921	0.003
Average Internal Rating	0.187	0.688	0.494	0.260

Table 1 – Correlation between ratings and *awvci*

We calculated *awvci* for window sizes of 1 and 5 days. With a time window of 1 day, contribution index values, which form the basis for *awvci*, fluctuate too much. A Window size of 5 days gives better results, smoothing peaks of activity and inactivity periods. It corresponds to a 5-day work-week and fits well into the overall project period of one month.

2.2 A ‘balanced’ internal communication leads to a better external rating

In this case the correlation between the *awvci* and the ratings is high (see Table 1). The external ratings show a higher correlation with the balance of the team’s communications than the internal ratings. It can be assumed that external ratings are more honest than the internal ones as students are not asked to rate team members they have been working with closely for a few weeks. They are more precise too, as they are based on a larger number of judgments.

2.3 Internal evaluation (peer) and external (by other group) ratings are correlated

The better the external team rating, the better the average internal rating of the team (Pearson Correlation=0.651; P-Value=0.113; n=7). A satisfied team gives good mutual ratings and provides work of good quality. This shows again that efficient teamwork has a positive impact on results (figure 4).

² There was no correlation between individual grade and contribution index neither.

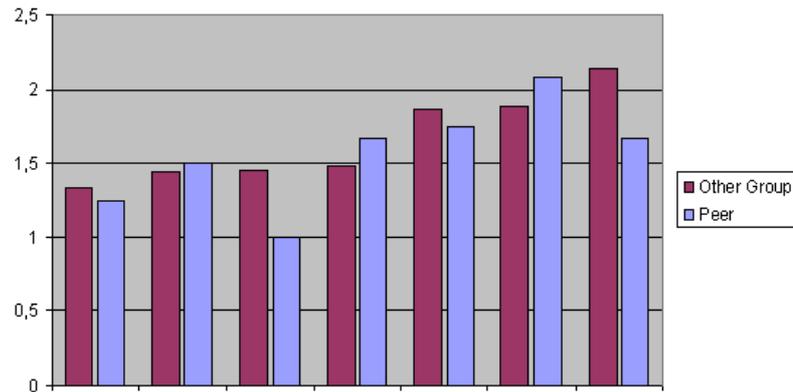


Figure 4 – Internal (peer) and external (by other groups) ratings of 7 teams (lower is better)

2.4 Limitations

While these early results are promising, they have to be taken with more than a grain of salt. The used dataset is small and somewhat incomplete. Communication was not completely recorded when it went through channels different from e-mail. Some teams sent messages to their team e-mail address to record these interactions, others did not. Ratings were done on a subjective basis with an underlying rigid structure. Also, our emphasis on temporal balance of contribution index only captures a subset of all communication activities.

3. CHALLENGES OF VIRTUAL COLLABORATION

The student groups faced several challenges during their virtual collaboration, which they reported at the end of the course. The students had not met each other face-to-face across countries, thus they did not know each other or their working styles, which caused some confusion and also getting a sense of “team work” was felt hard to achieve. The beginning was clearly the most difficult phase for many groups, it seemed to be quite hard to start an efficient work process and it took some time before a productive working mode was achieved.

The student groups were formed during a videoconference session: the students joined the groups according to their interest on suggested topics. The only rule for forming groups was that all the groups should have students from both countries. If a group had at least two students from the same country, this led to the formation of co-located sub-groups that at least partly communicated through other means than electronic (e.g., phone or face-to-face), thus this communication was not recorded and other group members could not follow it. Especially for groups having only a one-student sub-group in the other country, this caused difficulties for the isolated student to follow activities in the other country. Even though e-mail was the main communication medium, some groups started to use Skype (contains both chat and voice) or other chat programs especially for coordinating the work and making decisions. The synchronous communication was regarded as very efficient, but the problem was to find meeting dates suitable to all group members, since the students had many other courses at the same time. This problem often led to Skype or chat sessions between only two members at the time.

E-mail communication functioned quite well, but it was regarded as a less efficient communication medium than Skype or chat, since it was slow and thus not very interactive. Especially decision making was felt to be difficult through e-mail. Moreover, the asynchronicity of e-mail communication created uncertainty when

others did not know how to interpret the silence of the non-responding team member. Interpreting the sent e-mail messages was not always easy, neither. Translating from “Finnish” English to “German” English and vice versa opened up room for wide interpretation!

Despite these challenges the student groups did very good work and gained interesting results from the analysis of both the on-line communities and their own communication. The student feedback was very positive – the students felt that despite of the problems they had learned a lot.

4. LESSONS LEARNED

By organizing this course we learned a lot both regarding the arrangements of a distributed course and regarding data collection for research purposes. In the beginning of the course we did not give the students much advice on how to communicate or how to record the communication. We just asked the students to send a copy of all e-mails to an e-mail-box where all the communication of each group would be archived to be used when analyzing communication during their second assignment. We also offered MediaWiki as a forum for discussions. We learned that in the future it might be beneficial to teach in the beginning of the course some rules about how to work and communicate efficiently in a distributed team. In this course the students had to figure it out by themselves and make all the mistakes first, which of course took time away from working on the projects.

Since the student groups found Skype and chat very useful, we will need to encourage use of this kind of communication channel in the future. Moreover, a way to systematically record this kind of communication should be designed and taught to the students. Even though there was active communication inside the groups, the communication across the groups was very limited and took mainly place in connection with class videoconference sessions. Encouraging communication across groups will be needed, e.g. for solving technical problems. For instance, a discussion forum for technical communication problems could be started. Questions to the teachers could be directed to this forum, allowing everybody to follow and participate in these discussions.

The country-specific sub-groups were the reason that not all communication was recorded, e.g. phone calls and face-to-face conversations. This communication was often invisible to other team members, causing problems especially to one-person country-specific sub-groups, when he or she was to a certain extent left outside of the team. This posed additional problems for our communication research setup. This problem could be solved either by forming more balanced groups (at least two persons from one country), advising the students to record the non-electronic communication and informing the others, or by choosing only one team member from each site and organizing the course across several sites. This later solution is what we plan for the next version of this course: to involve four universities, which means four participating sites. That kind of a course would be both more challenging to organize and more challenging for the students to work in, when the groups are highly distributed. However, it would also be more interesting both for the students and for studying the communication patterns. Moreover, all the communication across these sites would be electronic and thus easier to record and for other team members to follow.

5. CONCLUSIONS

In this paper we presented our experiences of organizing a novel course on optimizing online communication behavior. The distributed student teams applied social network analysis to analyze communication behavior both in a chosen online forum and afterwards inside their own group. We obtained preliminary results on

correlating temporal online communication patterns with team performance. These results based on self-evaluation indicate that students in teams exhibiting balanced communication behavior performed best. Students used the insights they gained on the correlation of their own communication behavior with their group performance to improve their future communication behavior and collaboration style in COINs.

The presented communication analysis can only be considered indicative, as not all the communication was documented and as there were problems in the data recording. Despite these weaknesses, this experiment can be regarded as successful: the student feedback was very positive and we gained valuable ideas for both improving the course and better recording the communication data. Based on this experience we plan to teach this course again, this time among four universities, which will make the experiment both more insightful, but also more challenging! We would like to close with a quote from a student commenting on the course:

“This course was a great one. We learned a lot of things. The most valuable thing I learned was that the better communication is, the more successful you are (personally or as a team).”

6. ACKNOWLEDGEMENTS

The authors would like to thank the participants of this course from Cologne and Helsinki for being collaborative innovators themselves, coming up with many ideas for improvement of the course. Juha Itkonen and Kai Fischbach gave excellent suggestions based on an earlier version of this paper.

7. REFERENCES

- Cross, R., & Cummings, J.N. 2004, “Tie and network correlates of individual performance in knowledge-intensive work” *retrieved December 2004 at URL: http://ccs.mit.edu/fow/cross_cummings.pdf*
- Cummings, J., & Cross, R., 2003 “Structural properties of work groups and their consequences for performance” *Social Networks*, 25(3), 2003, 197-210
- Gloor, P., Laubacher, R., Dynes, S., Zhao, Y., 2003, “Visualization of Communication Patterns in Collaborative Innovation Networks: Analysis of some W3C working groups”. *Proc. ACM CKIM International Conference on Information and Knowledge Management, New Orleans, Nov 3-8, 2003.*
- Gloor, P. 2006, “Swarm Creativity, Competitive advantage through collaborative innovation networks” Oxford University Press, 2006
- Gloor, P., Zhao, Y., 2004, “A Temporal Communication Flow Visualizer for Social Networks Analysis”, *ACM CSCW Workshop on Social Networks. ACM CSCW Conference, Chicago, Nov. 6. 2004.*
- Kidane, Y., Gloor, P., Correlating Temporal Communication Patterns of the Eclipse Open Source Community with Performance and Creativity , NAACSOS Conference, June 26 - 28, Notre Dame IN, North American Association for Computational Social and Organizational Science, 2005
- Leenders, R.Th.A.J., Van Engelen, J.M.L., Kratzer, J., 2003 “Virtuality, Communication, and New Product Team Creativity: A Social Network Perspective”, *Journal of Engineering and Technology Management*, 20, 2003, pp. 69-92.
- Lueg, C., Fisher, D., 2003, “From Usenet to CoWebs, Interacting with Social Information Spaces”, *Springer, 2003.*
- Tyler, J., Wilkinson, D., Huberman, B. A., 2003 “E-mail as Spectroscopy: Automated Discovery of Community Structure within Organizations” *HP Laboratories, 2003. Retrieved February 2005 at URL <http://www.hpl.hp.com/shl/papers/e-mail/index.html>*
- Wasserman, S., Faust, K., 1994, “Social Network Analysis, Methods and Applications”, *Cambridge University Press. 1994.*