Exploring changes in network structures during online discussions

Tarja-Riitta Hurme, University of Oulu, Department of Educational Sciences and Teacher Education, Research Unit for Educational Technology, PO BOX 2000, Fin-90014 University of Oulu &
University of Turku, Centre for Learning Research, Educational Technology Unit, Fin-20014 University of Turku
Koen Veermans, Tuire Palonen, University of Turku, Centre for Learning Research, Educational Technology Unit, Fin-20014 University of Turku
Sanna Järvelä, University of Oulu, Department of Educational Sciences and Teacher Education, Research Unit for Educational Technology, PO BOX 2000, Fin-90014 University of Oulu
Email: tarja-riitta.hurme@oulu.fi, koen.veermans@utu.fi, tuire.palonen@utu.fi, sanna.jarvela@oulu.fi

Abstract: The aim of this study was to examine and model changes in communication network structures during online discussions in a virtual course at the university level. Two small groups of Finnish students (N=46) from different universities attended a national virtual course concerning a specific topic in the history of science. The course assignment consisted of participating into seven different discussion topics using Optima, an asynchronous learning environment. The discussions were used to examine communication networks during the course. Social network analysis techniques were used to explore communication networks and changes in these networks at structural and dyadic levels in four discrete time points with the effects of degree, reciprocity, transitivity and 3-cycles modeled by using the conditional method of moments. The results of the study suggest that during a virtual course the participants tend to form subgroups and hierarchical structures as in the same way as any other social community.

Aim and theoretical background

Research of online learning has focused on the organizational, technical, social and motivational factors that facilitate learning and more recently also on changes in interaction processes (Hewitt, 2005). One reason is to learn more about how and why the discussions die out, and if this can be attributed to the way that participants work in networked environments. It could be that they relay most of their attention on new notes, ignore older notes and after reading formulate one or more replies which indicates that they are using a single-pass strategy (e.g. Hewitt, 2005, 2003). This paper presents a social network analysis method using the Siena program to examine changes over time in communication networks at structural and dyadic levels during an online course. Siena software (Simulation Investigation for Empirical Network Analysis, v.3; Snijders, Steglich, Schweinberger & Huisman, 2007) is part of the StOCNET program suite (http://stat.gamma.rug.nl/stocnet/). Siena has been mainly used in examining the changes in friendship networks within the field social sciences (e.g. Van de Bunt, Duijn & Snijders, 1999; van Duijn, Zeggelink, Huisman, Stokman & Wasseur, 2003) but here it will be used to examine changes in communication structures in a networked learning environment.

In Hewitt's (2005, 2003) studies changes in networked discussions were examined with a computer simulation that mimicked the general behavior patterns of students' participation using exploratory modeling. Exploratory modeling is commonly used to examine how complex phenomena develop out of simple interactions in decentralized systems (Resnick & Wilensky, 1998). In Hewitt's studies, the computer simulation was used to predict how many computer notes could be included in one thread when the size of the threads varied. On the basis of the estimation results he suggested that single-pass strategies, which the students used, could be used to explain why the discussions die out in online learning (Hewitt, 2005, 2003). An alternative way to explore changes in networked interaction is to focus on the changes at structural and dyadic levels by using an actor-oriented model (Snijders, 2001, 2005). In the model, actor orientation means that the changes in an individual's connections to other participants can be dependant on the network structure and on attributes by observed covariates (Snijders, 2005).

There are several reasons why taking a social network perspective in examining changes in communication networks might unfold important information about network discussions. It gives for instance information about the strength of the connections in a communication network; connections that need to become strong enough before participants feel encouraged to take an extra effort to build and maintain a learning community. Second, the contact among participants needs to be supported by tutors to provide opportunities for the participants to share their knowledge and experiences in order to promote group interaction (cf. Haythorntwaite, 2000). It also gives information about reciprocal interaction, something that recent research in online learning has shown to be important for successful discussions. It means that in order to have successful

discussions, it is important that the participants do not only send their own explicitly formed thoughts to the discussion forum but also to reply to messages sent by others (Arvaja, Rausku-Puttonen, Häkkinen & Eteläpelto, 2003). There is evidence that the groups of students who are engaged in a collaborative knowledge building process ask questions and help to develop each other's thinking further. Reciprocal interaction has been considered to be a vitally important part of making the cognitive processes shared at a social level (Resnick, Levine & Teasley, 1993).

Method

Participants and working conditions

The participants of the study were 46 students (age 20-50) from six different universities throughout Finland. The participants were attending a specific course on the history of science. The participants were divided into two groups by a course organizer on the basis of their last names. During the course there were two main activities. Participants worked on writing an essay individually, and in addition to this task, they were required to participate in seven discussion topics in collaboration with their peers. The individual essay and discussions were not directly linked, but the students were encouraged to learn collaboratively and construct knowledge together. The course had three tutors, a male tutor who was responsible for technical guidance. Two female tutors were facilitators in group discussions; tutor 1 worked with group B and tutor 2 worked with group A.

The discussions took place in Optima, an asynchronous learning environment. The learning environment contained altogether nine discussion areas, two of them were general discussion areas in which the participants were able to ask questions about course requirements and ask for technical support. Further, there were seven separate group discussion areas, one area for one discussion topic for both groups. The first discussion forum consisted of the first week discussion where the participants told something about their scientific background and why they attended the course. During the following five weeks the students had five discussions on different themes concerning the specific issues in the history of science. These five discussion forum data were used in the analysis. In the seventh week discussion, the participants got feedback from the tutors and they evaluated their own performance in the course. Thus, for examining the patterns of interaction and for exploring the changes of communication networks in time, the five discussion topics concerning history of science were included to the analysis. These five discussion topics were thematically separated from each other.

Data collection

For the both groups, group A and group B, the analysis took into account the five discussion topics concerning the history of science. The discussion forum data consist of the participants' 852 posted computer notes; 496 notes were contributed by the members in group A and group B's discussion forum data consisted of 356 posted computer notes. In addition, the following attributive data of individuals were collected from the participants' messages during the first discussion. The individual attributes were *gender* (male, female), *position* in the virtual course (student, tutor), *university* (the university the student were attending), and *status* in the individual's university (master student, graduate student, researcher/lecturer, professor).

The discussion forum data was examined separately in the two small groups, group A and group B. The unit of the analysis was a computer note, a message contributed to the learning environment by a participant. The discussion threads were used in examining the patterns of interaction during the course. As a thread we mean a set of computer notes with the opening message which starts a new discussion topic and the computer notes written as a reply. The opening message was considered to be at level 1. The computer note that replied to the opening message was considered to be at level 2. Further, a message that replied to the level 2 computer note was situated to the level 3. The opening messages were not calculated in the communication matrices but they were taken into account when calculating the total amount of sent (outgoing) and received messages.

Data analysis

In the data analysis, the patterns of network interaction were examined by using multidimensional scaling (MDS) and by calculating basic social network analysis measures (Scott, 1991; Wasserman & Faust, 1994). The data was analysed with Ucinet (Borgatti, Everett & Freeman, 2002). In order to get a more detailed picture what kinds of changes there were in communication networks in time, the social network analysis with Siena program (Snijders, Steglich, Schweinberger & Huisman, 2007) were performed.

Multidimensional Scaling, MDS

The basic idea behind multidimensional scaling (MDS) is that of using the concepts of space and distance to map relational data. In this study, the relational data consists of the posted computer notes, the sent and received messages among participants. Multidimensional scaling is an attempt to convert graphic measures into metric measures (Wasserman and Faust, 1994; Scott, 1991). The goodness of each multidimensional scaling map can be measured by a value of stress, where a greater value means a poorer model. The value is, however, dependent on the data: the number of actors and the scales of measures. Analyses were performed on a symmetric matrix.

The symmetry was produced by summing up the received and sent messages. In this way both halves of the matrix became similar.

Social network analysis measures

The relational structures can be used to describe social phenomena where interactions between participants are observed. The patterns of interaction are examined first in order to clarify with whom the students worked and how many replies they sent and received. To obtain a more detailed picture of the features of interaction in the communication networks the following measures of the social network analysis (Scott, 1991; Wasserman & Faust, 1994) were used: *sent message, sent reply, received reply, sent to others, received from others* and a *betweenness* value indicating the position of the student pair in the communication network. A detailed description of the measures is given as follows.

1.) *Sent message* is the number of computer notes the participant sent to the other participants.
2.) *Sent reply* (out-degree value in directed asymmetric matrix) is the number of computer notes the participant sent as a reply to the other participants.
3.) *Received reply* (in-degree value in directed asymmetric matrix), shows the number of computer notes the other participants sent as a reply to the participant.
4.) *Sent to others* (out-degree value in a dichotomized matrix) is the number of other participants to whom the participant sent a reply.
5.) *Received from others* (in-degree value in a dichotomized matrix) shows from how many other participants the participant has received a reply.
6.) *Betweenness* shows how often a network actor's position is strategically important in the communication structure, i.e. how central the network position is. Interactions between two nonadjacent actors (i.e. actors who are not directly interacting) depend on the other actors who lie in the paths between the two. A participant has a high betweenness value if (s)he lies between two actors in the network who are not directly connected to each other, given that the shortest distance between two actors in the network is used to calculate the betweenness (Freeman, 1979; Wasserman & Faust, 1994).

Modelling the changes of communication network structure in time with Siena

In order to analyse the changes in network discussions during a virtual university course, the statistical estimation for the changes in social networks with the actor-oriented model was performed (Snijders 2001, 2005). Actor orientation means that an individual whose outgoing ties are changing in the network can depend on the network structure and on attributes by observed covariates. Thus, it is allowed that the probabilities of the relational changes depend on the entire network structure (Snijders, 2005).

The basic idea behind statistical modeling of social networks is that it uses the network structure as given at the first time point in order to model the changes in network structure. In addition, it is assumed that network processes are in dynamic state and there is unobserved network evolution going on between numbers of discrete time points (Snijders, 2005). In this study of the network structure, the first discussion is used as a starting point in modeling the changes at tie level in two different time points (time point 1 and time point2, time point 2 and time point 3, etc). Also the following measures *average degree, mutuality index* and *transitivity index* are calculated (Snijders, 2005). The mutuality index was defined as the fraction of ties reciprocated. The transitivity index is the number of transitive triplets divided by the number of potentially transitive triplets.

In the modeling, the objective function effects on the network variable were *outdegree, reciprocity, transitive triplets* and *3-cycles*. Transitive triplets are formed when for example Jaakko sends a message to Tuomo and Pirjo. Then Tuomo sends a message to Pirjo, but they don't have a reciprocal tie, meaning that although Tuomo sent a message to Pirjo she does not reply to the message. The 3-cycles are formed when Jaakko sends a message to Tuomo who sends the message to Pirjo, and she sends a message to Jaakko. In addition to these structural effects, the effects of four covariates, *gender* (male, female), *position* in the virtual course (student, tutor), *university* (the university the student were attending), and *status* in the university (master student, graduate student, researcher/lecturer, professor) were examined.

Results

In the results section, we focus on the patterns of interaction in communication networks formed by the participants during the online virtual course with the MDS-maps. Then we present the results of the social network analysis measures in order to take a deeper look how often and with how many others the participants were working with. After that we present the results concerning the changes in time in the communication networks.

MDS

The patterns of interaction among the two groups are displayed as MDS-maps in Figure 1. For the group A the stress value was .045 and group B's stress value was .024 which indicates a good model.

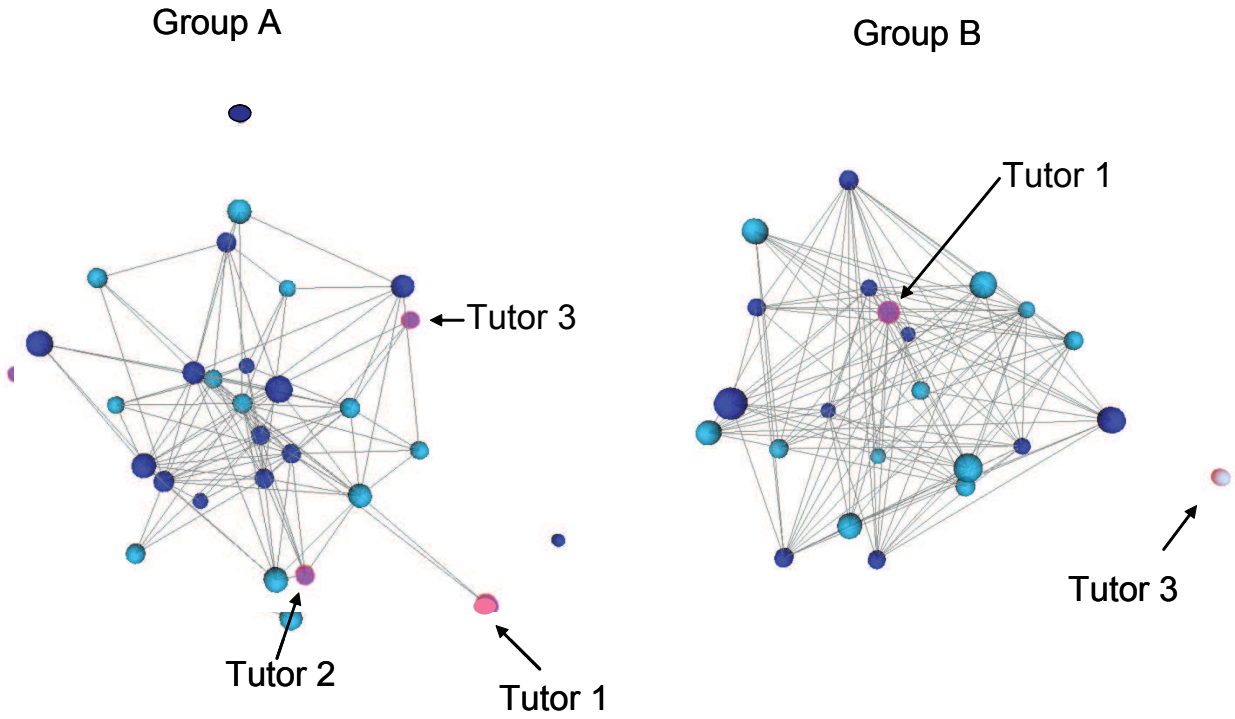

Figure 1. The communication networks in a virtual course.

In the figure 1, it can be seen that there were altogether three tutors, of which one tutor was responsible in technical issues (tutor 3) there a two other tutors were responsible in contextual tutoring. Tutors have been marked on the map, and other balls indicate students. In the group A, there are 3 tutors and 21 students, and in the group B there two tutors and 25 students. Comparing the MDS- maps, it seems that the group A's network is more loose that group B's. There is one isolated student in the communication network in group A. In Group B, the isolated participant is the tutor 3 who responsibility was to focus on tutoring in technical issues. He has sent messages to the discussion forum only if needed because in the Optima learning environment there was a specific discussion forum for the technical problems.

Participants' activity during the online course

The communication networks can be also examined with the social network analysis measures, for example sent messages and the amount of sent and received replies. Another interesting viewpoint is to explore how many other participants the students or tutors sent replies to, and from how many other participants the students or tutors received a reply. Also the betweenness value shows, how central the participant's position in the communication network is. The results of the participants' activity in online discussions are presented in the Table 1.

Table 1: Group's and tutors' activity during online course.

	Sent messages		Sent replies		Received replies		Sent to others		Received from others		Betweenness	
Participants	M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.
Group A	20,96	22,58	13,7	18,58	13,83	14,43	6,48	4,93	6,56	4,61	4,14	7,32
Tutor	4,67	4,73	4,67	4,73	3,66	3,79	3	2,65	3,5	1,53	0,03	0,05
Group B	16,38	16,87	13,24	13,85	12	13,33	7	3,79	6,43	4,23	3,28	7,31
Tutor	6	8,45	3	4,24	16	22,62	1,5	2,12	7,5	10,6	0,625	0,88

To get an overview, the average (M) and standard deviation (S.D.) of the social network analysis measures, were calculated for students and tutors in both groups (see Table 1). In general, the activity among participants varied a lot, although the participation in both groups was quite similar. In the communication networks, there were central participants whose betweenness values were high. There were also participants whose betwenneess value was close or equal to zero. These participants did not have a central position in the network. In addition, the tutors were not so active during the course, which can be seen in their sent message values being low. This can be explained by the students' way of working with tutors. It may be that they do not send so many replies to the tutor for clarification questions and thus it is likely that the tutors do not get replies. In addition, it may be that the content of the tutors' message were meant to facilitate the whole group, although it has been written as a reply to one student's message. In addition, in group A the tutor 2 had to suddenly give up the tutoring because of her other duties at the university, and the tutor 1 had to take a responsibility also for the group A. In the Table 2, the results of the basic social network analysis measures are used to provide a detailed picture of how the students considered as "stereotypical" cases participated into discussions.

Table: 2. Students' different profiles in communication network.

Group and Id	Sent messages (replies + opening comments)	Sent replies	Received replies	Sent to others	Received from others	Betweenness
A3	40	36	25	11	13	12,60
A5	35	15	43	7	16	13,05
A8	104	87	52	19	17	32,55
B3	16	16	1	10	1	0,47
B7	5	0	3	0	3	0
B20	31	31	15	11	3	1,15
B21	85	67	64	18	19	34,27

A student in group A with id 8, (later student A8) has been an active participant in the group's discussions and sent altogether 104 messages (Table 2). Student A8 has sent replies to 19 other participants of whom only two did not replied. In addition, student A8 has sent 87 replies and got 52 replies. This students' position in the communication network is central which can be seen in the high betweenness –value. Student A5 was active in starting the discussions and (s)he participated quite actively throughout the course, as was student A3, although (s)he did not sent as many opening comments to the discussion.

Student B21 sent messages to almost all other participants, only four other participants did not send any messages to him. Student B21 has the highest betweenness value indicating that (s)he had a strategically central position in the communication network. Student B3 is in opposite situation because (s)he sent only replies to the other participants and received replies from only one participant. It may be that student B3 is just sending as many messages as possible to make sure that (s)he passes the course. Another interesting case is student B20, who did not send any opening comments but (s)he sent, on average, three replies to 11 other participants and 5 recipients have sent 15 replies to him/her. Student B7's way of participating indicates a goal to pass the course with minimum work by starting 5 different threads after which (s)he had withdrawn from the discussions.

To sum up, these results showed that there were different kinds of participation among group A and group B. In group A, there were many students who sent opening comments and participated actively into the discussions. On the contrary, in group B, there were only few students who sent opening comments and maintained the discussion by sending replies to others.

Changes in time at structural and dyadic levels

In this section, a more detailed picture concerning the changes in communication networks will be provided. First, we present a general view of what kind of changes there were during the interaction at dyadic level. Secondly, we present the results concerning the ongoing ties (sent messages) after which the model of changes is provided. The dyadic level changes are presented in Table 1.

Table 3. An overview of the structure of networked discussion in groups A and B.

Time	T_0		T_1		T_2		T_3		T_4	
Group	A	B	A	B	A	B	A	B	A	B
Average degree	2.27	2.09	1.35	2.09	1.73	2.36	1.27	1.64	2.27	2.36
Mutuality index	0.40	0.17	0.15	0.29	0.23	0.29	0.23	0.17	0.43	0.26
Transivity index			0.07	0.4	0.28	0.27	0.09	0.05	0.16	0.28

The results in Table 3 shows that in group A and group B the average degree, starting at virtually nil, is changes from 1.3 up to a value of about 2.4 indicating that the participants sent in maximum three messages. The mutuality index is a quite low at almost 0.45. The transitivity index is also varying a lot, being quite low at 0.4. The time point 3 can be considered as a critical point in both groups because the average degree and transitivity index decrease when the discussion topic changes. It may be that the subject is not familiar with the students, which, in turn, decreased their activity. Another explanation could be that because the online discussions were not designed to support the individual course assignment, it may be that the participants had to focus on their essays before the deadline. The changes in ties between different time points are presented in Table 4.

Table 4: Change frequencies $N_{hk}(m)$ for the periods $t_m - t_{m+1}$.

Changes in ties, $N_{hk}(m)$	Period of observation $m \in [0,3]$							
	0 $(t_0 - t_1)$		1 $(t_1 - t_2)$		2 $(t_2 - t_3)$		3$(t_3 - t_4)$	
	A	B	A	B	A	B	A	B
0 to 0	620	430	630	426	633	429	619	426
0 to 1	**23**	**30**	**37**	**34**	**24**	**25**	**50**	**44**
1 to 0	**47**	**30**	**27**	**28**	**36**	**41**	**24**	**28**
1 to 1	12	16	8	18	9	11	9	8

The change counts (0 to 1, 1 to 0) are indicated in Table 2 Table 2 which shows that there are similar tendencies in the groups. The total number of changes between consecutive observation moments in group A is 70 and in group it is 60 in the first period. In the further periods the amount of changes varied from 60 to 74 and in group B from 62 to 72 in all further periods. At the first time point, in group A, there were 23 new communication ties formed and 47 ties broken, in group B there were 30 new and broken ties. Comparing the last time point, it can be seen that there are more new ties than broken ties. It may be that during the last discussion the students have been trying to increase their participation in discussions in order to pass the course. On the other hand, the topic may have been easier for them to participate in if they have been able to use their former knowledge and experiences in discussion.

To get a detailed picture what kinds of changes there were in group A's and group B's communication networks, we modeled with Siena how much the amount of sent message were different, was there more reciprocity during discussions and did the communication network have hierarchy indicating that there are participants who had a more strategically important position in the network than the others. The estimated results are presented in Table 5.

Table 5: Estimated changes in communication networks in group A and group B.

Estimated changes in communication networks	Period of observation $m \in [0,3]$							
	0 $(t_0 - t_1)$		1 $(t_1 - t_2)$		2 $(t_2 - t_3)$		3 $(t_3 - t_4)$	
	A	B	A	B	A	B	A	B
Outdegree	-2.57	-2.09	-2.38	-1.76	-3.03	-2.45	-1.98	-2.47
	(0.29)*	(0.21)*	(0.21)*	(0.19)*	(0.39)*	(0.57)*	(0.17)*	(0.48)*
Reciprocity	1.42	0.99	1.63	0.96	2.72	1.19	1.63	0.83
	(0.56)*	(0.44)*	(0.43)*	(0.37)*	(0.59)*	(0.41)*	(0.35)*	(0.38)*
Transitive	0.06	0.27	0.51	0.17	0.61	-0.00	0.43	0.02
triplets	(0.29)	(0.13)*	(0.12)*	(0.11)	(0.20)*	(0.21)	(0.16)*	(0.14)
3-cycles	-0.41	0.45	-0.90	0.00	1.63	0.24	-0.42	0.53
	(1.00)	(0.35)	(0.42)*	(0.29)	(0.70)*	(0.51)	(0.49)	(0.29)

*Statistically significant at p<.05 level.

The results in Table 5 show that the amount of sent messages varied a lot in different time periods. For both groups, the outdegree values are negative and statistically significant at .05 level. Further, reciprocity is positive and statistically significant ($p < .05$). These two estimates indicated that the tendency is not to make new communication ties but the participants continue the discussions with the same students, which, in turn, increased reciprocity.

In group A, the first two time point indicated that the amount of transitive triplets are increasing, which indicates that the discussion paths become longer and the closeness among participants increases. In other words the participants tended to participate in discussions with those from whom they received replies in different discussion areas. At the time points 2 and 3, the 3-cycle effect is statistically significant at .05 level. Thus it can be interpreted that there is a moment of hierarchy in the communication network. Thus, the participants' activity is unevenly distributed, and the same participants are active in many discussions at the same time. At other time points the participation is more evenly distributed and some active participants are involved in every discussion topic.

In the first two time periods, it can be seen that, in group B, the amount of transitive triplets increased. This can be considered as an increase in closeness within the communication network. In other words, the participants have a tendency to send comments to those participants from whom they had received replies in different discussions. The results also show that in group B there are not many 3-cycles which indicates that there is a hierarchy in the discussions. The participation is unevenly distributed among the participants and the same students are active in different phases of the online course.

In both groups, group A and group B, the estimation concerning individual attributes like gender, position, university and status, did not provide any statistically significant results. So, it can be said that in online courses, the participants do not choose their communication partners on the basis of individual attributes; for example, they do not get involved with students from their own university.

Conclusions

The aim of this study was to examine and model changes in communication network structures during online discussions in a virtual course at university level. In the analysis, social network analysis methods, multidimensional scaling and social network analysis measures were used to examine the patterns of interaction in two small groups. Further, social network analysis with Siena program was used to explore the changes in communication networks at structural and dyadic levels. The MDS-maps indicated that group B's network is more dense than group A's network. However, the social network analysis measures indicated that there are no big differences between the groups although there were more sent messages in group A than in group B. In group B there were less different discussions threads than in group A's discussion. A reason for the differences in the network structure between the groups can be found in the 3-cycle effect of Siena modeling. The structures were different because the 3-cycle effect in group A indicated that the same participants were active in many discussions at the same time. Thus, it can be said that the results of the study suggested that during a virtual course the participants tend to form subgroups and hierarchical structures in communication networks as in any other social community. In addition, the results show that participants did not choose their communication partners on the basis of individual attributes. Maybe something like that more extensive research on the individual attributes could help to find out whether these play a role in these interactions.

In this study, multidimensional scaling, social network analysis measures and social network analysis with Siena program was used to examine patterns of interactions and the changes in communication network in time. It can be argued that with these multiple methods we are able to get more insight what happens during the virtual courses. In addition, Siena program can be used to examine changes in communication structures in a networked learning environment. For future research, it would be important to study whether the quality of the discourse has an effect on changes in time in communication networks.

References

Arvaja, M., Rausku-Puttonen, H., Häkkinen, P. & Eteläpelto, A. (2003). Constructing knowledge through role play in a web-based learning environment. *Journal of Educational Computing Research* 46, 161-179.

Borgatti, S.P., Everett, M.G. and Freeman, L.C. (2002). Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

Freeman, L. C. (1979). Centrality in Social Networks: Conceptual Clarification. *Social Networks 1*, 215-239.

Haythorntwaite, C. (2000). Online personal networks: Size, composition and media use among distance learners. *New Media & Society,* 2(2), 195-226.

Hewitt, J. (2003). How habitual online practices affect the development of asynchronous discussion threads. *Journal of Educational Computing Research,* 28 (1), 31-45.

Hewitt, J. (2005). Toward an understanding of how threads die in asynchronous computer conferences. *The Journal of the Learning Sciences,* 14(4), 567-589.

Resnick, L. B., Levine, J. M., & Teasley, S. D. (1993). *Perspectives on socially shared cognition*. Washington: American Psychological Association.

Resnick, M. & Wilenky, U. (1998). Diving into complexity: Developing probabilistic thinking through role-play activities. *The Journal of the Learning Sciences*, 7, 153-172.

Scott, J. (1991). *Social network analysis: a handbook*. London: SAGE.

Snijders, T.A.B. (2001). The Statistical Evaluation of Social Network Dynamics. In M. E. Sobel and M. P. Becker (Eds.), *Sociological Methodology,* pp. 361-395. London: Basil Blackwell.

Snijders, T.A. B. (2005). Models for Longitudinal Network Data. In P. J. Carrington, J. Scott, and S.Wasserman (Eds.) *Models and Methods in Social Network Analysis*. New York: Cambridge University Press.

Snijders, T.A.B., Steglich, C.E.G., Schweinberger, M. & Huisman, M. (2007). Manual for SIENA version 3. University of Groningen: ICS, Department of Sociology.

van Duijn, M.A.J., Zeggelink, E.P.H., Huisman, M., Stokman, F. N. & Wasseur, F.W. (2003). Evolution of sociology freshmen into frienship network. *The Journal of Mathematical Sociology,* 27(2), 153-191.

Van de Bunt, G.G., Duijn, M.A.J. & Snijders, T. (1999). Friendship Networks Through Time: An Actor-Oriented Dynamic Statistical Network Model. *Computational & Mathematical Organization Theory,* 5(2), 167–192.

Wasserman, S., & Faust, K. (1994). *Social network analysis: methods and applications*. Cambridge: Cambridge University Press.