# Two make a network: using graphs to assess the quality of collaboration of dyads

Irene-Angelica Chounta[1], Tobias Hecking[2], H. Ulrich Hoppe[2], Nikolaos Avouris[1]

[1] HCI Group, University of Patras, Greece
{houren, avouris}@upatras.gr
[2] Collide, University of Duisburg-Essen, Germany
{hecking, hoppe}@collide.info

**Abstract.** In this paper we explore the application of network analysis techniques in order to analyze synchronous collaborative activities of dyads. The collaborative activities are represented and visualized as networks. We argue that the characteristics and properties of the networks reflect the quality of collaboration and therefore can support the analysis of collaborative activities in an automated way. To support this argument we studied the collaborative practice of 228 dyads based on graphs. The properties of each graph were evaluated in comparison to ratings of collaboration quality as assessed by human experts. The activities were also examined with respect to the solution quality. The paper presents the method and the findings of the study.

**Keywords:** learning analytics·collaboration·analysis·network·graph theory·SNA·CSCL

## 1 Introduction

The analysis of collaborative activities is a popular research area in the interdisciplinary fields of CSCL (Computer-Supported Collaborative Learning) and CSCW (Computer-Supported Cooperative Work). It allows researchers to gain insight into the nature of collaboration, to support designers and orchestrators and to provide feedback to users with respect to their practice. CSCL constitutes an ideal application field for automated analysis techniques based on activity log files. Interaction analysis and automated metrics of interaction have been widely used for the quality assessment of collaborative activities [1, 2]. Several activity metrics, such as the number or rate of messages, the average word count, the roles alternations and the symmetry of actions, have been introduced as automated methods of analysis [3, 4]. However, the nature of data sets limitations to that kind of methods. Data-driven approaches are not stemming from a theoretical framework and it is argued that the lack of qualitative research undermines the depth of analysis and entails drawbacks [5, 6].

There are several examples of using graph representations and Social Network Analysis (SNA) analysis techniques for the analysis of learning activities [7, 8]. Most of the studies focus on individual students or groups of students of bigger size in order

adfa, p. 1, 2011.

to reveal patterns of interaction and minimize the complexity of the analysis [9]. Graph theory has also been used to describe and analyze cognitive structures during collaborative learning activities [10]. In this paper, we propose the use of networks for the quality assessment of collaborative activities that involve dyads of students. Even though dyads are common in collaborative learning scenarios, they are hardly considered as “networks” or studied with SNA techniques. We argue that the interaction of dyads can be effectively represented as a network and that the network properties will reflect the quality of the collaboration. In order to support this argument, we used the log files of collaborative activities of dyads to construct networks that represent interaction structures. The graphs were analyzed and studied in comparison to the ratings of collaboration quality as assessed by human evaluators. A correlation analysis between the network properties and the ratings of the evaluators was conducted.

The paper is organized as follows: In the second section we provide an overview of related work, in section 3 we present the background of the study, the collaborative setting and the dataset used. In section 4 we describe the method of the study. The results are presented in section 5 and the paper concludes to a discussion on the findings as well as future work.

## 2 Related Work

### 2.1 Generating networks from event logs

The representation of log files as networks has been part of existing and ongoing research on collaboration analysis. An advantage is the explicit representation of relations between entities that are not directly observable by looking at log files. Networks can be further analyzed and modeled by using well established techniques of graph theory [11].

A well known framework for the graph-based representation of joint activities is described in [12]. Based on tables of event logs, relations between events are extracted according to predefined contingency relations. These contingency graphs can be further processed to extract graph based models for interaction, mediation and sociograms. Nasirifard et al. [13] use subject-verb-object representations of read and write actions on documents in a shared workspace to extract multi-relational networks between users and objects, as well as relations between the users themselves.

One problem that arises with the network representation of activity logs is that one loses an explicit representation of time. While logs usually carry a timestamp, the network representation aggregates log based relations over a certain period of time. One solution can be to sample networks from action logs in subsequent time windows. This results in subsequent time slices of an evolving network. This technique has been used in [14] in order to analyze changes of the affiliations of groups of students to learning resources over time. In some cases however, networks bear an inherent notion of time. This is the case when the nodes of the network relate to events, and directed links between nodes establish a partial temporal ordering, as in citation networks [15]. The networks described later in this paper also have this property of an implicit notion of time. Consequently the applied analysis does not require sampling the networks into time slices.

### 2.2 Theory of network evolution

It could be shown that the evolution of many real world networks does not follow a random model. Moreover, there are fundamental mechanisms that govern the emergence of the given network structure. Hence, the structural characteristics that can be observed by investigating those networks are potentially meaningful. This is the basic argument to apply network analysis methods in order to reveal basic patterns in network data. In this study, networks are constructed from event logs of collaborating dyads and the assumption is that those networks reflect the quality of the collaboration. This section reviews the major research on networks models that helps to interpret the results presented in section 5. The Erdös Rényi model [16] describes the evolution of graphs when the likelihood of a connection between any two nodes is uniformly distributed. In those networks, the distribution of the degree of the nodes (number of connections) follows a Poison distribution where most nodes have a degree close to the average degree. The average path length in such networks tends to decrease when new edges are added to the network [17]. However, it can be shown that this network generation model does not apply to most real world networks. Moreover, the degree distribution often follows an inverse power law instead of a Poisson distribution [18]. This means that the majority of nodes is not well connected but, however, the probability to have a few highly connected nodes (hubs) is much greater than in random networks according to the Erdös Rényi model. These networks are scale-free because of the power law degree distribution. Although most nodes in scale-free networks are sparsely connected with the rest of the network, there are densely connected regions (clusters). The hubs often bridge between those clusters. One model that explains the emergence of scale-free networks is the preferential attachment model [18]. In this model, nodes connect more likely with nodes that already have many connections.

Recently it could be shown that scale-free networks can also emerge when nodes are subsequently added to a network. In this model new arriving nodes connect to existing nodes by chance but also to nodes that their neighbors already connect with [19]. This model comes quite close to networks that are constructed from subsequent events, as in this work.

## 3 Background of the study

In this study we explore the analysis of collaborative activities through networks. We use a dataset of collaborative activities that has been previously evaluated with qualitative and quantitative methods involving human judgment, so it can serve as a reference and benchmark. The dataset consists of 228 collaborative sessions that took place during a programming course in the University of Patras (Greece). In particular, dyads of students were asked to collaborate synchronously through a groupware application in order to complete a task. The task was the collaborative construction of an algorithmic flowchart. The duration of the activity was about 90 minutes. The activity was mediated by a shared workspace groupware application [20]. The application comprises two shared spaces: (a) a common workspace for the construction of diagrammatic representations and (b) a chat tool to support the synchronous communication of users (Fig. 1). The dataset has been used in earlier studies by human experts for the empirical evaluation of a rating scheme. In addition, it was further analyzed for the training of a machine learning model for the classification of collaborative activities. The findings of earlier studies were used as valuable input in the current case.

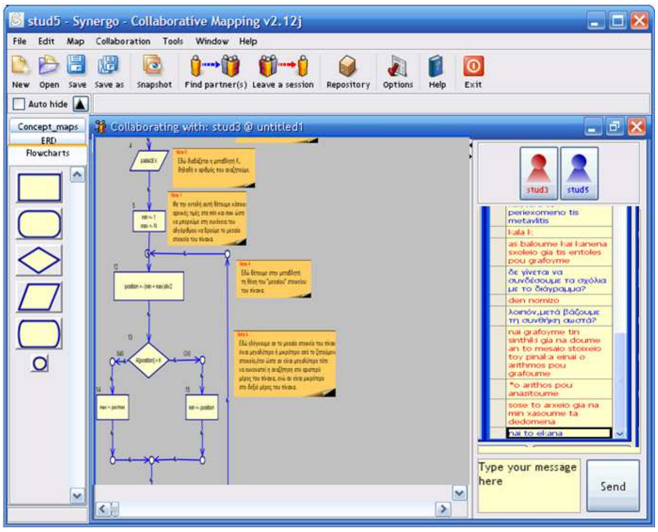


**Fig. 1.** The user interface of the collaborative application Synergo. It consists of the common workspace used for the construction of diagrammatic representations and the chat tool that mediates user communication.

### 3.1 Assessment of collaboration quality by human experts

The dataset was previously evaluated by two human experts with the use of a rating scheme [21]. The evaluators rated each collaborative session on various aspects of collaboration with respect to their quality with the support of a rating handbook and after a training phase. According to the scheme (Table 1), the quality of collaboration was rated on six dimensions that represent the four fundamental collaborative aspects: Communication, Joint information processing, Coordination and Interpersonal relationship. Each dimension was rated on a 5-point Likert scale ([-2, +2]). The general dimension of the quality of collaboration (Collaboration Quality Average, CQA) was computed as the average of all six collaborative dimensions. The ratings were successfully tested for inter-rater reliability and consistency. Depending on the collaborative dimension, ICC ranged between 0.83 and 0.95, while Cronbach's alpha ranged between 0.91 and 0.98. Therefore, the rating scheme was proved to be a successful means for the assessment of the quality of collaborative activities.

**Table 1.** The rating scheme used for the assessment of the quality of collaboration: the collaborative dimensions in relation to the general aspect of collaboration that they represent

| General aspects of collaboration | Collaborative Dimensions |
|---|---|
| Communication | Collaboration flow |
| | Sustaining mutual understanding |
| Joint information processing | Knowledge exchange |
| | Argumentation |
| Coordination | Structuring the problem solving process |
| Interpersonal relationship | Cooperative orientation |

### 3.2 Assessment of collaboration quality with the use of time series

The current dataset was further analyzed and evaluated using time series analysis techniques [22]. In particular, the collaborative activities were represented as time series of aggregated events within certain time frames. A memory-based learning model was used for the classification of collaborative sessions with respect to the similarity of time series. The research hypothesis was that collaborative activities of similar collaboration quality will unfold in similar ways.

The memory based model was used as an automatic rater of the quality of collaboration (CQA). The ratings of the model were compared to the ratings of human experts. We conducted a correlation and error analyses in order to evaluate the efficiency and accuracy of the model. The analysis showed that the ratings of the model correlated with the ratings of expert evaluators ($\rho=0.3$, $p<0.05$), while the mean absolute prediction error MAE was less than one on a 5-point scale (MAE=0.89). It was also shown that time series of activity aggregated within small time frames (60 seconds or less) portray collaboration more effectively. The effect that the size of time frames has on the performance of the model was further explored. It was shown that the time frames ranging from 15 to 30 seconds are the most suitable for depicting the quality of collaboration, thus pinpointing that meaningful interaction for synchronous communication takes place within time frames of less than a minute. This finding comes into agreement with other related studies [1, 23].

We use the results of previous studies to reveal potential relations between the network properties and the ratings of human experts on collaborative dimensions. The research question we aim to explore is whether the quality of collaboration is depicted on the network that represents the practice of the dyad and how this information can be used to support the analysis of collaborative learning activities.

## 4 Method of the study

In the present study, we aim to assess collaboration quality using network analysis. To that end, each collaborative session is represented as a network and is visualized as a graph. The nodes stand for the actions taking place in the common workspace and the chat tool. The edges of the network represent some kind of dependency between actions or relation. Therefore, in order to consider two, or more, actions as relevant and thus, connect them, the following conditions should apply:

(a) The time distance between actions must fall within a certain range. Based on the findings of previous study [22], the time range signifying relevant actions is defined from 10 to 30 seconds.
(b) Workspace actions are considered relevant, and therefore connect, only if they involve the same artifact. In the case of chat messages, only temporal proximity is taken into consideration.
(c) The identity of the actor should differ. The objective of the study is to map the interplay between users rather than the practice of a single user. Therefore we map consequent actions of different actors, aiming to reveal *reciprocal activity*.

The instance of a log file used for the construction of the networks is portrayed in Fig. 2. The log files are structured according to the OCAF format [24] as:

*<ID><timestamp><actor><event-type><attributes>*

where:

- <ID> is an incremental identifier, unique for each event
- <timestamp> the time that the event occurred
- <actor> the user responsible for the event
- <event-type> the type of the event, i.e. chat message or the type of the workspace action
- <attributes> a field related to additional information that can be interpreted in combination with the type of event, i.e. content of chat message or (x-y) coordinates of an artifact on the workspace.

Taking into consideration the aforementioned rules for constructing nodes and edges, we elaborate on the paradigm of Fig. 2. For the log file entries with ID number from 18 to 22, there will be five nodes, one for each entry. The node "18" will be connected through an edge with node "19" because these actions come from different actors and their time distance is less than 30 seconds. Likewise, the node "19" will be connected with nodes "20" and "21" but not with node "22" since their time distance is more than 30 seconds. The instance of the network that corresponds to the instance of the log file of Fig. 2 is presented in Fig. 3.

| | | | | |
|---|---|---|---|---|
| 17 | 0:07:51 | ece7108 | Set object to front | [Alternate process (1)] |
| 18 | 0:07:46 | ece7108 | Change Concept Entity text | [Alternate process (1), Start, (x=385,y=32,w=90,h=60)] |
| 19 | 0:08:02 | a08-6930 | Chat message | [then we will write the contents of A and B] |
| 20 | 0:08:12 | ece7108 | Chat message | [right] |
| 21 | 0:08:28 | ece7108 | Chat message | [it says so in the book] |
| 22 | 0:08:43 | ece7108 | Chat message | [that we should use this shape] |
| 23 | 0:08:51 | ece7108 | Insert Entity | Input-Output, (x=382,y=127,w=90,h=60), Input-Output (2) |

**Fig. 2.** Instance of the log file of a collaborative activity as recorded by Synergo.

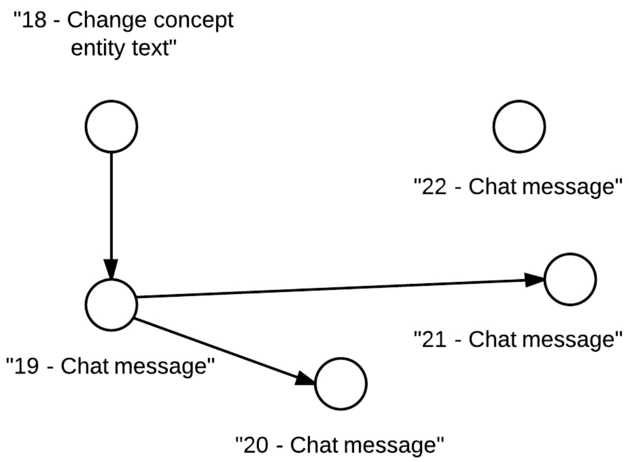


**Fig. 3.** Example of a network constructed by the log file instance of Fig. 2

The graphs of collaborative activities were constructed by the web-based application SiSOB Workbench [25]. Two networks, one for a collaborative session of good quality and one session of poor collaboration quality respectively, are portrayed in Fig. 4. On one hand, the session of good collaboration quality is portrayed as a big network that depicts intense activity, i.e. a large number of user actions. A number of nodes appear to be strongly connected to each other revealing some kind of activity bursts caused by key events while the rest are sparsely connected. On the other hand, the session of poor collaboration quality is portrayed as a smaller network. The instances of activity bursts are less and not so tightly structured as in the previous case.

We argue that the quality of collaboration is reflected by the characteristics of the network that represents each activity. To support this argument, we study the corresponding basic network properties and compare these to the ratings of collaboration quality, as described in section 3.1. A correlation analysis is carried out in order to establish whether the network properties correlate to the quality of collaboration.

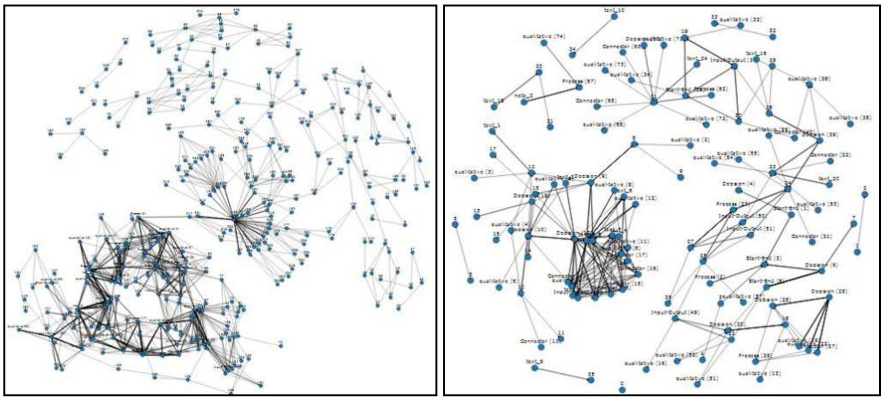

**Fig. 4.** Two networks representing a session that was rated as of good quality of collaboration (left) and a session of bad collaboration quality (right)

The network properties used in the study are:

- The number of nodes (N)
- The number of edges (E)
- The density of the network (D). The density of a network is defined as the ratio of the number of edges to the number of possible edges: $= \frac{2E}{N(N-1)}$
- The diameter of the network (d). The diameter of a network is defined as the longest of all the calculated shortest paths in a network
- The average path length (apl). The average path length is the ratio of the sum of all shortest paths between all pairs of nodes to the sum of total number of pairs. It is used to indicate the number of steps needed to move from one node to another
- The clustering coefficient average (cc_avg). The clustering coefficient is defined as the ratio of existing links connecting a node's neighbors to each other to the maximum possible number of such links. We use the global definition of the clustering coefficient for the network as a whole. This is defined as the number of closed triangles (three fully connected nodes) and open triangles (three nodes and two pairs are connected)
- The assortativity (A). Assortativity is described as the tendency of nodes to connect with other nodes, similar to themselves
- Power Law Fits (plFits). To verify if the network can be considered as a scale-free network, a power law has fitted to the degree distribution of the networks. Then a Kolmogorov-Smirnov test has been conducted in order to measure the goodness of fit. The smaller this value is, the better is the fit of the power law distribution.

## 5 Results

In the current study, we used the dataset of 228 collaborative activities, as described in section 3. Each one of the activities was represented as a network by applying the guidelines as described in section 4. The aforementioned network metrics and properties were computed for each collaborative session. A correlation analysis was carried out between the network properties and the ratings of human evaluators on the collaboration quality average (CQA). The network properties were also studied with respect to the individual collaborative dimensions as defined by the rating scheme and also the grades assigned by the teachers for the solution quality. Spearman's correlation coefficient was used since the data are not normally distributed.

### 5.1 Network properties and collaboration quality

The results of the correlation analysis are presented in Table 2. All the correlations portrayed in Table 2 are statistically significant ($p<0.05$). The network metrics of clustering coefficient and assortativity were correlated neither with the collaboration quality average (CQA) nor with the solution grade on a significant level. The correlations that were not statistically significant are not presented in the results. Most of the network properties that correlate to the quality of collaboration also correlate to the individual collaborative dimensions.

**Table 2.** Correlations of the network properties and the ratings of collaboration quality average (CQA) and for each collaborative dimension of the rating scheme

| **Collaborative Dimensions** | **#Nodes (N)** | **#Edges (E)** | **Density (D)** | **Diameter (d)** | **(plFits)** | **(apl)** |
|---|---|---|---|---|---|---|
| Collaboration flow | 0.358 | | -0.405 | 0.267 | -0.229 | 0.226 |
| Sustaining mutual understanding | 0.351 | 0.179 | -0.275 | 0.246 | | 0.205 |
| Knowledge exchange | 0.400 | 0.190 | -0.340 | 0.228 | -0.192 | 0.191 |
| Argumentation | 0.416 | 0.169 | -0.360 | 0.273 | -0.225 | 0.223 |
| Structuring the problem solving process | 0.339 | 0.156 | -0.268 | 0.185 | -0.208 | 0.154 |
| Cooperative Orientation | 0.410 | 0.195 | -0.323 | 0.295 | -0.196 | 0.257 |
| CQA | 0.446 | 0.180 | -0.394 | 0.294 | -0.233 | 0.243 |

The number of nodes correlates highly with the quality of collaboration ($\rho_N$=0.446). On the other hand, the number of edges correlates on a low level with collaboration quality ($\rho_E$=0.180). This shows that sessions with intense activity are also of better collaboration quality. However, the number of connections does not necessarily point to better quality. The number of nodes correlates significantly with all collaborative dimensions ratings and in particular with those that stand for the aspect of Joint information processing (Knowledge exchange and Argumentation). These dimensions assess how effectively students exchange information and give appropriate explanations to their partners. Thus, it is expected that a student who is motivated to explain and to make a valid argument will post a higher number of explanatory messages or add helpful notes/artifacts on the common workspace, resulting in intense activity.

In the case of good collaboration quality networks, there is a significant negative correlation of the collaboration quality (CQA) with the value of the Kolmogorov-Smirnov test for the power law degree distribution, where a small value indicates better fit ($\rho_{plFits}$ = -0.233). Hence, the networks generated from good collaborative sessions appear to be scale-free networks. These networks have some nodes that are strongly connected with other nodes while the majority is only sparsely connected. Since the nodes of the network represent user activity and they can only link to events that happened later in time (condition [a]), one can conclude that good collaborations contain a few key events that cause bursts of subsequent events related to the key event. The relation of activity bursts and the scale-free property has been studied thoroughly [26].

The diameter (d) and the average path length (apl) as metrics of the linear size of a network also correlate positively with the quality of collaboration ($\rho_d$=0.294, $\rho_{apl}$=0.243). This finding indicates that good collaboration usually results in big scale-free networks with some key events. Those networks are characterized by long paths that correspond to longer uptake chains. The longer path indicates that the teams that collaborate efficiently are more influenced by previous events than those who collaborate poorly, since condition [b] requires connected events to be on the same object. Therefore, collaborative activities of good quality appear to have more key events that dominate user activity and direct future focus. Taking into account the temporal condition for the network construction, it is evident that activities of good collaboration quality unfold faster in time. The diameter and average path length correlate with the ratings of human experts on all collaborative dimensions, especially on the collaborative aspect of Cooperative orientation.

Density (D) is a metric to measure the saturation of relationships between the nodes of a network in relation to a completely connected network. One would expect that good collaboration would lead to denser networks. However the results indicate that dense networks stem from collaborative practices of bad quality ($\rho_D$ = -0.394). Consequently, density is a negative predictor of collaboration quality in the sense that dense networks point to poor collaboration. An explanation for this counterintuitive finding is that in scale-free networks the density is anti-proportional to the number of nodes. Therefore, a higher density indicates a smaller map [27]. In addition, density correlates negatively to all individual collaborative dimensions. The highest negative correlation is scored for the dimension of Collaboration flow, which stands for the aspect of Communication ($\rho$ = -0.405). Collaboration flow, in particular, expresses how naturally communication flows between actors and among shared resources. It is therefore expected that good communication flow is depicted by a bigger network map and thus, lower density.

### 5.2 Network properties and solution quality

Apart from the log files of the activity, the students had to deliver the final algorithmic diagram, as it was developed in the shared workspace. The diagram was graded separately by the teacher of the course on a scale from 0 to 10. The assessments of the solution were straightforward since the task had a demonstrably correct solution with a few easily defined alternatives in parts of the diagram.

The correlations are presented in Table 3. The properties that do not appear statistically significant correlations are not included. The solution grade correlates significantly with the number of nodes and edges of the networks. Lower positive correlations also appear for the solution grade and network the diameter and average path length. The results of the correlation analysis show that thorough solutions, i.e. detailed flowcharts that result in intense activity and, therefore, bigger networks, are evaluated as good. This is expected since a correct solution requires a fully developed flowchart, accompanying notes etc.

We should note that the solution grade does not correlate either to the network density or the power law degree value. As aforementioned, these network properties characterize collaborative practice and meaningful interplay among users. In this case however, a good collaboration does not ensure a good solution quality and vice versa. For example, a detailed solution could be the result of one student’s work while his partner does not contribute on the task. This would be an example of a good solution grade and bad collaboration quality. On the other hand, a good collaboration does not ensure that the students will succeed in their task. This would result in a bad solution grade but good collaboration quality. The difference between solution quality and collaboration quality is depicted in the results of the correlation analysis and the network properties.

**Table 3.** Correlations of the network properties and the ratings on solution quality

| | #Nodes (N) | #Edges (E) | Diameter (d) | Average Path Length (apl) |
|---|---|---|---|---|
| solution grade | 0.319 | 0.305 | 0.202 | 0.189 |

## 6 Conclusions and Future Work

This study explores the use of networks for the analysis and evaluation of collaborative activities of student dyads. Graphs and social network analysis techniques have been introduced for the analysis of the collaborative practice of small or bigger teams that usually work together over discussion forums or supported by various groupware applications [27, 28]. Even though dyads are a popular way of students grouping in a learning context, they are hardly considered as networks. We argue that the application of networks for the mapping of dyads interaction, can indicate the meaningful interplay and successful collaborative practice, as well as offer a valuable tool for the assessment of collaboration quality.

For the purpose of the study, we analyzed a dataset of 228 collaborative sessions where each session was represented by a network. The network properties of each session were studied in combination to ratings of human experts on collaboration quality and a correlation analysis was carried out. The results showed that the quality of collaboration reflects certain properties of the networks that represent joint activities. The size of the network, in terms of number of nodes and diameter, points to activities of good collaboration quality. Efficient collaboration is expected to result in intense activity and therefore, bigger networks. It was also shown that good practices unfold faster in time and form longer chains of actions (i.e. longer paths). The size of the network is a good indicator for the solution quality as well. Collaborative activities represented by big networks are assessed with higher grades for the quality of the solution. We should note however that a good solution does not presuppose effective collaboration. This was also portrayed by the network properties. Collaboration quality was found to correlate negatively to the density of the network as well as the power law fits criterion value, while that was not the case for the solution quality. Density is used to measure the relationships of a network and the power law fits is used to indicate scale-free networks. The negative correlation shows that good collaboration results in scale-free networks where a key action leads to reciprocal interplay among the actors. This is portrayed in the form of few central nodes that are strongly connected to others while the majority of nodes is sparsely connected leading to low density and a small value for the power law fit criterion. In the case however that a good solution grade is the result of one-sided activity, i.e. one student who takes the lead, this kind of interaction is absent.

In this study, the construction of the networks was based solely on the temporal and spatial proximity of user activity. In future studies, we plan to use content analysis techniques in order to refine the relations and connections between user actions. This is believed to lead to more accurate representations of collaborative interaction and add up semantic value to activity patterns. The application of advanced methods such as path analysis and sequential data analysis is also an interesting direction that could be pursued in order to define patterns that indicate good or bad collaboration quality.

## References


1. Schümmer, T., Strijbos, J. W., Berkel, T.: A new direction for log file analysis in CSCL: Experiences with a spatio-temporal metric. In: 2005 Conference on Computer Supported Collaborative Learning (CSCL'05), International Society of the Learning Sciences, pp. 567-576 (2005)
2. Martínez-Monés, A., Harrer, A., Dimitriadis, Y.: An interaction-aware design process for the integration of interaction analysis into mainstream CSCL practices. In Analyzing Interactions in CSCL. Vol. pp. 269-291. Springer, (2011)
3. Bratitsis, T., Dimitracopoulou, A., Martínez-Monés, A., Marcos, J. A., Dimitriadis, Y.: Supporting members of a learning community using Interaction Analysis tools: The example of the Kaleidoscope NoE scientific network. In: Advanced Learning

Technologies, 2008. ICALT'08. Eighth IEEE International Conference on, IEEE, pp. 809-813 (2008)

4. Kahrimanis, G., Chounta, I. A., Avouris, N.: Study of correlations between logfile-based metrics of interaction and the quality of synchronous collaboration. In: 9th International Conference on the Design of Cooperative Systems, Workshop on Analysing the quality of collaboration, International Reports on Socio-Informatics (IRSI), Aix en Provence, pp. 24 (2010)
5. Stahl, G.: Rediscovering CSCL. In: CSCL 2: Carrying forward the conversation, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 169-181 (2002)
6. Stahl, G., Koschmann, T., Suthers, D.: Computer-supported collaborative learning: An historical perspective. Cambridge handbook of the learning sciences (2006)
7. Reffay, C., Teplovs, C., Blondel, F.-M.: Productive re-use of CSCL data and analytic tools to provide a new perspective on group cohesion. Connecting computer-supported collaborative learning to policy and practice 846-850 (2011)
8. Toikkanen, T.,Lipponen, L.: The applicability of social network analysis to the study of networked learning. Interactive Learning Environments 19, 365-379 (2011)
9. Palonen, T.,Hakkarainen, K.: Patterns of interaction in computersupported learning: A social network analysis. In: Fourth International Conference of the Learning Sciences, pp. 334-339 (2013)
10. Ifenthaler, D., Masduki, I., Seel, N. M.: The mystery of cognitive structure and how we can detect it: tracking the development of cognitive structures over time. Instructional Science 39, 41-61 (2011)
11. Wasserman, S.: Social network analysis: Methods and applications. Cambridge university press, (1994)
12. Suthers, D. D.: Interaction, mediation, and ties: An analytic hierarchy for socio-technical systems. In: System Sciences (HICSS), 2011 44th Hawaii International Conference on, IEEE, pp. 1-10 (2011)
13. Nasirifard, P., Peristeras, V., Hayes, C., Decker, S.: Extracting and utilizing social networks from log files of shared workspaces. In Leveraging Knowledge for Innovation in Collaborative Networks. Vol. pp. 643-650. Springer, (2009)
14. Hecking, T., Ziebarth, S., Hoppe, H. U.: Analysis of dynamic resource access patterns in a blended learning course. In: Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, ACM, Indianapolis, Indiana, pp. 173-182 (2014)
15. Halatchliyski, I., Hecking, T., Göhnert, T., Hoppe, H. U.: Analyzing the flow of ideas and profiles of contributors in an open learning community. In: Proceedings of the Third International Conference on Learning Analytics and Knowledge, ACM, pp. 66-74 (2013)
16. Erdős, P.,Rényi, A.: On the evolution of random graphs. Magyar Tud. Akad. Mat. Kutató Int. Közl 5, 17-61 (1960)
17. Watts, D. J.,Strogatz, S. H.: Collective dynamics of ‘small-world’networks. nature 393, 440-442 (1998)
18. Barabási, A.-L.,Albert, R.: Emergence of scaling in random networks. science 286, 509-512 (1999)
19. Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tomkins, A., Upfal, E.: Stochastic models for the web graph. In: Foundations of Computer Science, 2000. Proceedings. 41st Annual Symposium on, IEEE, pp. 57-65 (2000)
20. Avouris, N., Margaritis, M., Komis, V.: Modelling Interaction during small-group Synchronous problem solving activities: the Synergo approach. In: 2nd International Workshop on Designing Computational Models of Collaborative Learning Interaction,

ITS2004, 7th Conference on Intelligent Tutoring Systems, Maceio, Brasil, pp. 13-18 (2004)

21. Kahrimanis, G., Meier, A., Chounta, I.-A., Voyiatzaki, E., Spada, H., Rummel, N. et al.: Assessing collaboration quality in synchronous CSCL problem-solving activities: Adaptation and empirical evaluation of a rating scheme. In Learning in the Synergy of Multiple Disciplines. Vol. 5794, pp. 267-272. Springer, (2009)
22. Chounta, I.-A.,Avouris, N.: Time Series Analysis of Collaborative Activities. In Collaboration and Technology. Herskovic V, Hoppe HU, Jansen M, andZiegler J (eds.), Vol. 7493, pp. 145-152. Springer Berlin Heidelberg, (2012)
23. Suthers, D. D., Dwyer, N., Medina, R., Vatrapu, R.: A framework for conceptualizing, representing, and analyzing distributed interaction. International Journal of Computer-Supported Collaborative Learning 5, 5-42 (2010)
24. Avouris, N. M., Dimitracopoulou, A., Komis, V., Fidas, C.: OCAF: an object-oriented model of analysis of collaborative problem solving. In: Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community, International Society of the Learning Sciences, Boulder, Colorado, pp. 92-101 (2002)
25. Göhnert, T., Harrer, A., Hecking, T., Hoppe, H. U.: A workbench to construct and re-use network analysis workflows: concept, implementation, and example case. In: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ACM, pp. 1464-1466 (2013)
26. Barabasi, A.-L.: The origin of bursts and heavy tails in human dynamics. nature 435, 207-211 (2005)
27. Hoppe, H. U., Engler, J., Weinbrenner, S.: The Impact of Structural Characteristics of Concept Maps on Automatic Quality Measurement. In: International Conference of the Learning Sciences (ICLS 2012), Sydney, Australia, pp. (2012)
28. Reihaneh, R., Takaffoli, M., Zaïane., O., R.: Analyzing participation of students in online courses using social network analysis techniques. In: Proceedings of educational data mining. (2011)