T. Koschmann, D. Suthers, & T.W. Chan (Eds.), *Computer Supported Collaborative Learning 2005: The Next 10 Years!*
Mahwah, NJ: Lawrence Erlbaum Associates.

A New Direction for Log File Analysis in CSCL: Experiences with a Spatio-temporal Metric

Till Schümmer[1], Jan-Willem Strijbos[2], Thomas Berkel[1]

[1]Computer Science Department
FernUniversität Hagen, Germany

[2]Department of Educational Sciences
Leiden University, The Netherlands

Corresponding author: Till.Schuemmer@fernuni-hagen.de

Abstract. This paper discusses the importance and difficulties of assessing interaction between students. To ease the detection of interaction in student groups, a metric is developed that can measure the level of interaction based on log file data. The metric is based on a spatial model and detects actions that take place in close spatial or temporal proximity. After providing a formal definition of the metric, an exploratory analysis of interaction in two different settings is reported to determine the feasibility of the measure: synchronous interaction in a collaborative puzzle game and asynchronous interaction in student groups that use the BSCW shared workspace system.

Keywords: Log file analysis, Spatial models, Interaction awareness, Interaction assessment

INTRODUCTION

Collaborative learning has become a popular approach at most educational levels and increasingly so in higher education (Strijbos, Kirschner, & Martens, 2004). In the past decade most institutions in higher education have implemented Virtual Learning Environments (VLE's) (De Graaff, De Laat, & Scheltinga, 2004). Mostly these systems include a package of standard tools, such as a calendar, document sharing, a discussion forum, and a chat. Essential for such collaborative systems is 'interaction awareness', which help students to maintain an overview of their collaborative processes and provides the teacher a means for assessment.

Most VLE's support interaction awareness through generic notifications (e.g., an indicator signaling new documents) or e-mail digests with recent changes (Chyng, Steinfeld, & Pfaff, 2000). These mechanisms inform students and teachers when artifacts were changed and who made the changes, but they do not provide information about the degree of interactivity within the collaboration (i.e., who responded to whom – or made changes to an artifact - and how close were they related in time). For example, KnowledgeForum© includes an Analytical Tool Kit (ATK) providing descriptive information (e.g., on number of notes written) (Chan & van Aalst, 2004). In technical terms, changes to a calendar, document repository, or discussion forum are considered as manipulations on objects in the environment.

In CSCW research the evaluation of activities is considered as a key factor for improving the design of groupware sytems as it provides an overview of system use. Information about user activities is often recorded by means of log files that include an entry for each interaction of the user with the groupware system (examples have been collected by Pinelle and Gutwin (2000)). Log files usually contain a very large amount of activity data which needs to be transformed to activity reports to compile fine-grained activity data to large-grained indicators for assessment. In CSCL research such activity reports were initially used to assess the extent of collaboration, such as the number of messages (Harasim, 1993), mean number of words (Benbunan-Fich & Hiltz, 1999), thread-length (Hewitt, 2003), and 'social network analysis' (SNA; Lipponen, Rahikainen, Lallimo, & Hakkarianen, 2003). When analyzing log files, it is easy to determine if a user was active, but it is difficult to find out to what extent these activities contributed to the group process. A high level of activity does not imply that the actor is contributing to the group. This is often a wrong assumption read from log files. It is now widely acknowledged that activity reports provide a surface analysis of collaboration at best (Stahl, 2001) and most researchers have turned to in-depth studies of the communicative process (Strijbos, Martens, Prins, & Jochems, in press; Schümmer & Haake, in press).

These in-depth small scale studies are typical for CSCL and provide valuable insights in how knowledge is collaboratively constructed (Stahl, 2004), but they offer little consolation for teachers whose higher education institute has implemented a VLE and are subsequently confronted with large scale supervision and assessment requirements. Students' experiences with various teaching and learning environments are reflected in their study approach and preference (Entwistle & Tait, 1990) and there is growing evidence suggesting that students tune their learning activity to the assessment that is conducted (Scouller, 1997). Thus, if interaction and collaboration are integral parts of the didactical goals but are not assessed, they will not take place to the desired extent. Efficient means for assessing interaction are thus required in a didactical approach that focuses on interaction.

Equipped with activity reports only, temporal and spatial proximity (i.e., changes or addition closely related in time or position in the collaboration space) cannot be detected. Including temporal and spatial parameters can reveal patterns that can support students in their evolving collaboration. Such patterns can assist teachers in making inferences about collaboration efficiency and detect the need to intervene in poorly collaborating groups. Moreover, the generation of activity patterns based on temporal and spatial proximity can provide a better estimate of actual interactivity using log file data. Hence they can be implemented as group and teacher support in those settings where online activities occur on a large scale and where teachers can no longer monitor and supervise all groups in detail. One clue for determining the impact of interactivity can be a measure of success. If the interaction leads to success, it is less likely that the interactive work was not conflicting. However, measuring success is in general very difficult; especially in the case where work or learning concentrates on so-called "wicked problems" (i.e., problems with no fixed or right solution for example 'school dropout') (See Conklin & Weil, 1997).

In HCI research, log files have been widely used and automated tools exist for calculating metric information from the log files in single user applications, which provides clues on how the system is used by the individual user (for an overview on automated evaluation see Ivory & Hearst, 2001). Some metrics like the task completion time or the number of activities per time are easy to adapt and calculate in collaborative applications. An example for a group metric is shown by Begole, Tang, Smith and Yankelovich (2002) who detected the times where a user was actively using the computer and calculated an average activity rhythm chart for specific users or groups of users. This example shows how a single user metric (activities per time) can be transformed into a group metric. But still, the metric does not provide information on the interaction between group members, nor does it reflect that an individuals' work rhythm can be affected by the rhythms of other users.

In general, one can observe a lack in groupware specific metrics that provide clues on groupware-mediated interaction between users. One reason, why groupware specific metrics are rare could be the larger complexity of groupware settings. While one can often clearly define the different usage sequences in single user applications (e.g., the use of a pull-down menu that can invoke a specific action), collaborative applications have to deal with more than one control flow. For instance, discussion boards allow parallel postings of different users or graphical editors allow the concurrent creation of diagram elements. This means that the evaluation of multi-user log data needs to consider the individual users' interaction flows as well as the group interaction, which evolves from the users' actions. In this paper it is argued that such metrics can be calculated and complement analysis techniques known from single user applications. A specific group interaction metric is proposed that measures the degree of interaction in a group, based on temporal and spatial proximity. This metric can be used to extend the expressiveness of group rhythms (e.g., by inferring isolated and interactive parts of group work) for the evaluation of CSCW systems and it can provide clues for supervision and assessment in CSCL environments.

Before the details of the metric are discussed, it is essential to elaborate the difference between system feedback (such as in single user applications) and interaction feedback (in multi-user applications). Next, the mathematical model for the calculation of the proposed group interaction metric is discussed, followed by two examples to illustrate the feasibility of the metric in two distinct groupware systems: a collaborative puzzle game and the shared workspace system BSCW. In the final sections both examples will be contrasted and directions for future applications will be discussed.

A METRIC FOR MEASURING GROUP INTERACTION

The goal of the following analysis is to measure the degree of interaction between the users in a collaborative environment. Interaction between users can be defined as a set of two or more actions that mutually or reciprocally influence one another. In the context of collaborative applications, this means that users modify shared objects and other users adapt their activities according to activities perceived before. Since all interaction is computer-mediated, it can all be reduced to the process of modifying and perceiving shared objects. Close interaction means that users work on the same or on related objects (for example artifacts) at near points in time or in the collaborative space. An object is any information unit shown to the user.

Interaction feedback (IF) relates to three kinds of feedback in VLEs as shown in figure 1. Based on his mental model of the system (*"people form internal, mental models of themselves and of the things with which they are interacting; these models provide predictive and explanatory power for understanding the interaction"* (Gentner & Stevens, 1983)), user A performs object manipulations (1. OM_A). The system answers the manipulation with system feedback (2. SF_A). This can for instance be the update of a visual representation on the screen. At the same time, user B receives activity feedback of user A's activity (3. AF_A). While perceiving the modified object state, B changes his mental model (4. CMM) of the set of shared objects according to OM_A. The changed mental model may trigger an object manipulation of B (5. OM_B). As it was the case for A's object manipulation, OM_B triggers system feedback for B (6. SF_B) and activity feedback for A (7. AF_B). But since AF_B semantically replies to OM_A, it is interpreted as interaction feedback for A (IF_A). This kind of interaction feedback is relevant for detecting group interaction.

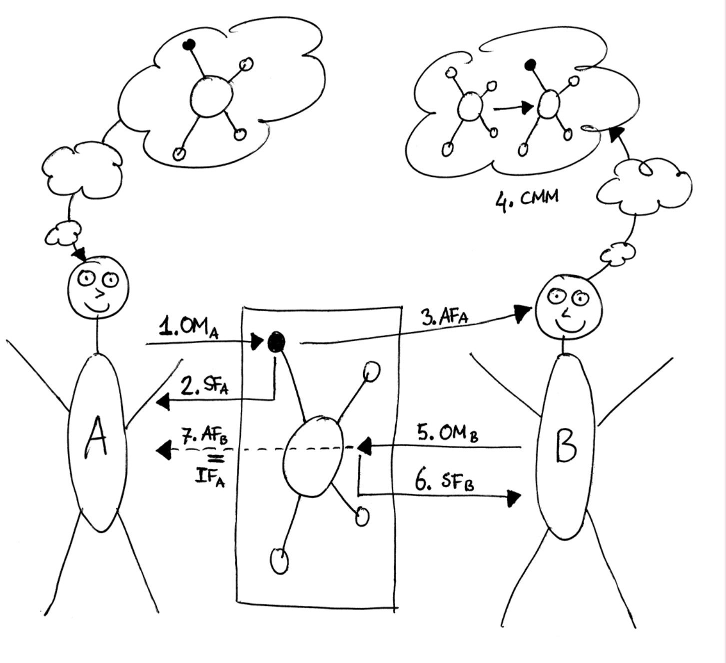

Figure 1: Schematic representation of different classes of feedback.

The difference between system or activity feedback and interaction feedback is similar to that between effects *of* technology (system feedback about manipulations) and effects *with* technology (manipulations as interaction feedback) (Salomon, Perkins, & Globerson, 1991). Naturally, the system generates both types of feedback simultaneously during collaboration, yet system feedback is explicit whereas interaction remains implicit (one could even say 'in the eye of the beholder'). Current activity reports of VLEs focus on system feedback and activity feedback but ignore interaction feedback in their calculations.

Instead of activity reports, group interaction can be calculated (and represented) using a metric based on the spatial awareness model (Rodden, 1996), which has been widely applied in synchronous groupware systems (cf. the *active neighbors* design pattern (Schümmer, 2004) for an in depth discussion of the awareness model and more known uses of and experiences with the model). The spatial awareness model consists of objects, such as artifacts and users, who are distributed in space. Each object has a well-defined distance to all other objects. The distance should reflect the semantic nearness between the two objects. In simple models it can be defined by distances of the objects' locations in a document storage. In this case, one assumes that a user will group related documents in related places of the document storage (e.g., the same folder). A more complex model could analyze the artifacts' content and compare its semantic (cf. Leximancer, a system for creating document spaces (Smith, 2000)). Another approach could be to analyze existing relations between artifacts in case that they are connected by hyperlinks. One example for this approach is provided in (Schümmer, 2002) where items of a web shop are related regarding to their descriptions.

As in the spatial awareness model, the strength of interaction between two activities is defined as the spatial distance of the manipulated objects. In cases, where the system consists of different functional components (e.g., a calendar that manages appointments, a discussion forum that stores contributions, a chat communication channel transferring chat entries, or a shared whiteboard containing graphical objects), it can be appropriate to calculate interaction only for actions on artifacts of the same functional component. Otherwise, a distance function between artifacts of different components is needed. It is assumed that strong interaction occurs when two activities are performed on two artifacts that have a low spatial distance. Besides spatial distances, interaction has to take a temporal dimension into account. It is assumed that two activities occurring at the same point in time – or closely related in time – imply a stronger interaction than two activities that occur at different points in time. Combining spatial and temporal relations of activities surpasses the unrelated information provided in single or multi-user activity reports. In the next section the mathematical calculation of the group interaction metric, termed '**I**nter**A**ction value' (**IA**), will be discussed in more detail.

Calculating the Interaction Value IA

A set of activities $a_1,...,a_n$ is considered as input for the mathematical calculation of the interaction value. Each activity has to provide information on the manipulated artifact, the time $t(a)$ when the activity a occurred, and the user $u(a)$ who performed the activity.

Since the spatial model requires a method for distance calculation between two activities, a function $ds(a_i,a_j)$ is defined that calculates the distance between the artifacts that were the focus of the activities a_i and a_j. In case of a spatial arrangement of the artifacts like in a shared drawing tool, this can be the difference between the positions of the touched diagram elements. In case of structural arrangements like folders in a shared file system, it can be the length of the path between the touched files.

The spatial distance $ds(a_i, a_j)$ has to be combined with the temporal distance $dt(a_i, a_j) = t(a_i) - t(a_j)$ between the two activities. Since space and time are two different dimensions, these are arranged in a two-dimensional vector space. The distance between the activities a_i and a_j can then be calculated as the Euclidean norm of the time distance and the space distance. The interaction $ia(a_i, a_j)$ between two activities a_i and a_j is formalized as

$$\mathrm{ia}(a_i, a_j) = \left| \begin{pmatrix} ndt(a_i, a_j, dt_{max}) \\ nds(a_i, a_j, ds_{max}) \end{pmatrix} \right|$$

with the normalized time distance (*ndt*)

$$ndt(a_i, a_j, dt_{max}) = \max\left\{0, \frac{dt_{max} - dt(a_i, a_j)}{dt_{max}}\right\}$$

and the normalized space distance (*nds*)

$$nds(a_i, a_j, ds_{max}) = \max\left\{0, \frac{ds_{max} - ds(a_i, a_j)}{ds_{max}}\right\}$$

The values dt_{max} and ds_{max} are upper bounds that define which distance activities are considered as relevant. The normalization of the time and space distance by dt_{max} and ds_{max} ensures that the measure for $ia(a_i, a_j)$ can be applied to different types of groupware applications. For example, when analyzing a synchronous groupware application a small value for dt_{max} will be chosen. To calibrate the metric, the values for dt_{max} and ds_{max} have to be found on an experimental basis. Too small values will always result in very small interaction values (a flat line with values close to 0), while too large values will always produce interaction values that are close to an upper bound. An appropriate choice for dt_{max} and ds_{max} will generate interaction values that are distributed between 0 and an upper bound.

It is also important to consider group size in relation to time. As group size increases each group member will have less opportunity to touch or manipulate objects within a given time-span, as users compete for time. This implies that within a large group, there is a higher probability that there will be more inactive people – such as lurkers who are merely following the discussion (in contrast to free-riders who have no intention to compete for time). Logging reading activities includes lurkers as active users who contribute to the group interaction value. Filtering reading activities on the other hand shifts the focus of *ia* to users who collaboratively construct content. Depending on the underlying learning approach – cognitive versus experiential learning as proposed by Rogers and Freiberg (1994) – reading activities can be considered as important or less important. Since experiential learning focuses on the active construction of a solution, modifying activities are central for judging successful group interaction in this case while reading activities can be filtered or rated as less important.

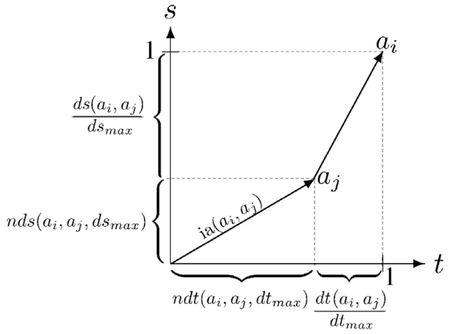

Figure 2: Calculation of the interaction value for two activities a_i and a_j

Figure 2 illustrates that the vector defining the distances between the two activities is subtracted from the vector (dt_{max}, ds_{max}). The result is normalized with respect to (dt_{max}, ds_{max}). Thus, activities that are near in space and time will result in an interaction vector that is close to the unit vector. The interaction value $ia(a_i, a_j)$ is then calculated as the length of the interaction vector.

Up to now, it has been shown how the interaction between two activities can be calculated. The equations can be extended to calculate the interaction of an activity with respect to all previous activities performed by other users within the bounds ds_{max} and dt_{max}. The rationale behind this is that interaction implies that a user reacts to the activities of other users.

For this accumulated interaction value, the sum of all activities' interaction values is calculated for those activities that were performed by other users. In order to calculate the interaction value at any point in time, the time distance calculation *ndt* must be adopted, so that it calculates the time distance between a point in time *t* and the average of the two activities' times $((t(a_i) + t(a_j)) / 2)$. The result is divided by the number *m* of activity pairs (a_i, a_j), which produced an interaction value $ia(a_i, a_j) > 0$. This leads to the following equations for the accumulated interaction value *IA(t)* for a fixed point in time *t*:

$$\mathrm{IA}(t) = \frac{\sum\limits_{\substack{a_i, a_j \in A: \\ t \geq t(a_i) \geq t(a_j) \wedge \\ u(a_i) \neq u(a_j)}} \left| \begin{pmatrix} ndt'(a_i, a_j, t, dt_{max}) \\ nds(a_i, a_j, ds_{max}) \end{pmatrix} \right|}{m}$$

with

$$ndt'(a_i, a_j, t, dt_{max}) = \max\left\{0, \frac{dt_{max} - \left(t - \frac{t(a_i)+t(a_j)}{2}\right)}{dt_{max}}\right\}$$

In the next section two examples are discussed that illustrate the calculation of the interaction value, as well as the interpretation of the calculation that can provide clues for students and/or teachers for adapting their collaboration or supervision practices.

TWO CASE STUDIES

The theoretical model was applied to log data from two different applications: a synchronous collaborative learning object (puzzle game) and the BSCW shared workspace system. The next two sections explain these applications and present our findings.

The COAST-Puzzle

A jigsaw puzzle game was chosen for three reasons. First of all, jigsaw puzzles are easy enough to understand for validating the proposed interaction calculation. Users collaborate in a spatial setting and the collaboration strength differs depending on the division of the puzzle space between the users (i.e., users can be active in the same part of the puzzle or work independently in different regions). Secondly, they have their roots in the early 19th century in educational settings (for an overview see Hannas, 1981). Nowadays, jigsaw puzzle games are mainly used in recreational contexts, but depending on the content of the puzzle it can support didactic goals (e.g., a jigsaw puzzle with outlines of countries that have to be assembled to form the European map). Finally, the degree of success (or progress towards the solution) can be measured at any point in time, which allows for comparing the calculated interaction value to the degree of success.

The game has simple rules: a picture of an animal is presented to the user for a short period of time (5 seconds), cut into 30 pieces, and finally scrambled. The players' task is to restore the original image by moving pieces with their mouse. Whenever a user moves a piece, it is moved on the other machines as well. To indicate, who moves the piece, a hand icon is placed on the center of this particular piece. In our example the puzzle was played by student groups of three randomly assigned players. All groups were co-located as shown in Figure 3 and the system stored all moves in a log file.

Figure 3: A group playing the puzzle game.

To apply the interaction value metric, spatial and temporal distances between the activities had to be calculated. The spatial distance was calculated as the minimum of the distances between the two start positions and the two end positions of the activities. Spatial distance was calculated by normalising the positions so that all pieces were squares and thus vertical and horizontal distances had the same impact. The time distance was calculated from the two activities' start times. The values for dt_{max} and ds_{max} were found on an experimental basis. Again, too small values produced too flat curves and too large values produced curves that were constantly at a high level. Finally, *IA(t)* could be calculated for the puzzle logs. Figure 4 shows the result of the *IA(t)* calculation (the thick curve) for which dt_{max} = 21.5 sec and ds_{max} = 3.0, were used, corresponding to the distance of three puzzle pieces.

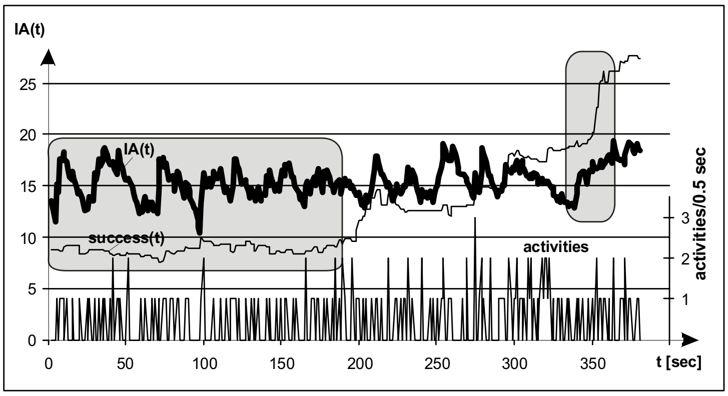

Figure 4: The interaction value *IA(t)* of a puzzle session.

In addition to *IA(t)*, Figure 4 also shows the success of the complete puzzle game as the thin curve. Succes was calculated as the accumulated differences of the pieces current mutual distances to their correct mutual distances. A low value in the success curve thus stands for higher entropy in the puzzle game, while high values represent more order. The third curve (bottom of Figure 4) shows the number of moves per time tick of approximately 0.5 seconds. This example of *IA(t)* reveals important characteristics of the metric that are needed to interpret the graphs. Whenever interaction takes place, the curve rises steeply. If nothing would happen after the interaction, the curve would continuously decline. Any raise in the declining curve indicates that actions relating to the initial pair of interactive actions took place.

By comparing the different curves, conclusions can be drawn regarding the degree of interaction while solving the cooperative puzzle game:

- There are time shifts visible between the curves. This expresses that users sometimes need more time to adapt their mental representation after they percieved another user's activity feedback and perform the required object manipulation. In the right box, this is visible as a peak of interaction, which later on evolves into a strategy of success accompanied by an above-average interaction value.
- A rise in the interaction curve does not imply that the success curve will rise as well. This is visible in the left box in Figure 4 where the interaction curve is variable and the success curve is constant.
- Changes in *IA(t)* provide clues for relevant interactive parts of the log file, which should be examined in more detail. This clue is qualitatively different from clues found in a quantitative analysis of actions per time frame (e.g., activity reports). Even if the size of the time frame is enlarged, the action per time frame curve does not provide sufficient information for detecting interactive sequences in the puzzle.
- The difference between *IA(t)* and the clues drawn from simply counting activities per time becomes clear by comparing the thick curve for *IA(t)* and the thin curve for the number of activities. A large number of activities do not imply a close interaction. Simply counting activities does not consider the relations between different activities.

It can be argued that the cooperative puzzle game represents an artificial interaction situation, hence the interaction value analysis technique was applied to the BSCW environment as this resembles more closely a component of any Virtual Learning Environment: document sharing.

BSCW

BSCW (Basic Support for Cooperative Work) provides folder based workspaces in a tree structure to support cooperative document sharing (Bentley et al., 1997). The main difference to the puzzle game is that users work asynchronously, while in the puzzle synchronous interaction took place. The system has been used in different courses at the FernUniversität in Hagen for several years now.

BSCW log files of a practical training on databases were examined that was held in the winter term 2002/2003. It involved seven groups with six to seven students working for a period of 102 days. Every group had its own workspace and no access to workspaces of other groups to prohibit lurking in foreign workspaces.

To support information exchange across groups, a special discussion workspace was added. Milestones for parts of the project work were defined by the project leaders.

During the course, every access on BSCW objects, such as documents, folders or message boards was logged. The mapping of the BSCW workspace structure to the spatial model of the proposed metric is based on the assumption that users will group semantically related artifacts in the same or related folders in the workspace. A spatial distance can then be calculated as the length of the shortest path between touched artifacts in the folder tree. Temporal distances can be calculated directly from the timestamps of the log entries. The upper bounds for the calculation of the *IA* value were dt_{max}=5.25 days and ds_{max}=3.0. Compared to the puzzle example, dt_{max} was extended to incorporate the asynchronous nature of BSCW. Fig. 5 shows the result for *IA(t)* for two groups *G1* and *G2*. Again, *IA(t)* is shown as a thick line. The second curve shows the number of activities per time sample.

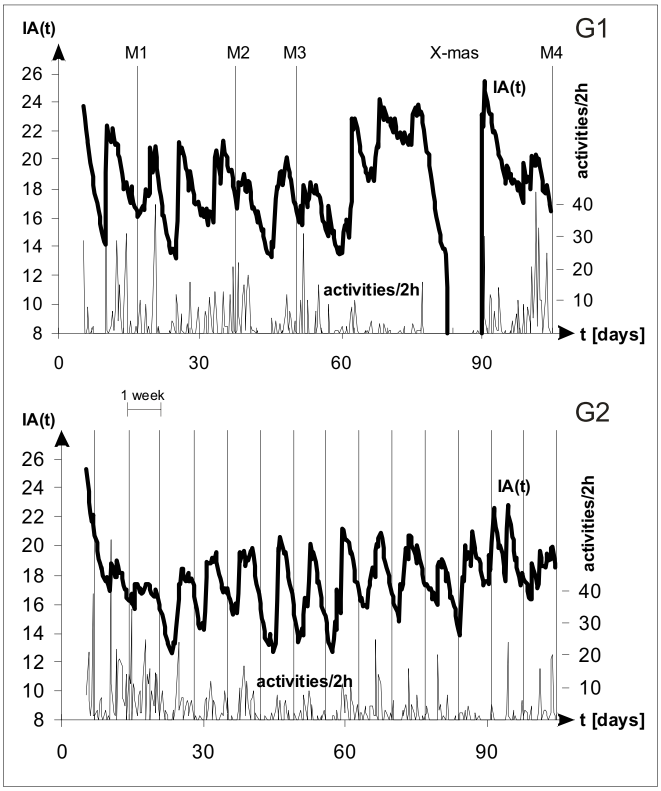

Figure 5: The interaction value *IA(t)* of two groups *G1* and *G2* detected in the BSCW log.

G1 started with a first peak in interaction that was needed to align the goals and create an exposé. This peak corresponds to the maximum before M1 in figure 5. The exposé was the first milestone determined by the teachers. This and three other milestones are shown in the diagram for *G1* as thin vertical lines. After the exposé was done, the group showed a short peak of interaction after the discussion with the course leader. The group had an interaction peak every other week, where they worked on the second milestone (M2: writing a project proposal). The interaction decreased after the proposal was done. In this period of time, the group was asked to refine the proposal and deliver this as M3. The collaboration on this task was postponed until shortly before the deadline. In the final phase, the group had to build a database system based on the project proposal. From the interaction value it can be inferred that the group started to collaborate on the topic about two weeks after M3. Actually, the combination of *IA(t)* with the curves for the number of activities reveals an interesting finding during this time frame: the users started with peaks in the activity curve that had accompanying peaks in the *IA(t)* curve (at a time of 60 days). But in the following days, the numer of activities decreased to a low value while the *IA(t)* value remained high, indicating that users collaborated on very close artifacts. The opposite relation can be observed close to M2. In this part of the diagram, the number of activities was relatively high but at the same time the *IA(t)* value declined. This indicates a situation where users interacted with the system at the same time, but did not relate their activities to activity feedback percieved of other users' activities. In other words, the users divided the task and worked on those parts individually. The first major peak of *IA(t)* after M3 indicates that the group members were realligning their activities.

During the holiday season (Christmas and New Year), the group did not interact at all. This long period of inactivity led to interaction values of 0. After the group started interacting again, the *IA(t)* curve shows a very large peak. This peak does not imply that the group interacted closely. Instead, it is an attribute of the mathematical model: all interaction value curves start with an initial peak since the calculation requires that several user actions can be set into relation. In the interpretation of the diagram, these initial peaks should thus be ignored (or understood as the start-off of group interaction). One last peak can be seen shortly before the final mile stone (M4). From the interaction curve, one can observe that there was no stable group rhythm. Instead, the group interaction peaked before the milestones M2 and M3.

G2 showed a different interaction style that reveals a quite stable group rhythm. Again, the peak in the first week is a result of the mathematic characteristics of the measure (since the measure needs some initial data to tune). The group started with no obvious interaction pattern in the second and third week. The first relevant part of the diagram is at the beginning of the fourth week, where group members worked mainly independent in the BSCW system, which resulted in a decline of *IA(t)*. After the fourth week, the group found its rhythm: they started to have a very regular alternation between interaction and independent work within a period of one week (the vertical lines in figure 5 represent one-week time spans). This rhythm remained stable throughout the course, even in phases with a low activity rate (during the holiday season).

Both interpretations of the log files match with the observations made by the course organizers during chats held at every mile stone.

DISCUSSION

In this paper, a new model for calculating interaction between group members was introduced. In contrast to the 'interaction' provided by activity reports, the alternative model is based on a distinction between three types of feedback resulting from object manipulations by users. Whereas current metrics are based on system and activity feedback, the proposed group interaction metric explicitly incorporates interaction feedback. It thus focusses on object manipulations that are executed in response to activity feedback of other users' actions. By necessity the theoretical model is slightly more extensive than the actual metric being calculated, as it also includes the perception of a change by another participant (awareness) which thrives interaction. Yet, including awareness in the analysis would require far more than log file data.

As shown by two examples, this model can be used to calculate a group interaction metric that incorporates temporal and spatial proximity. Although the puzzle game and the BSCW system have many differences, the same analysis model (defined as a function of temporal and spatial manipulation on objects in the electronic environments) can be applied to detect interaction: Regarding the application of the **I**nter**A**ction value (**IA**) four aspects appear to be relevant:

- **Synchronous versus asynchronous groupware:** This difference is compensated by using different values for dt_{max}. At least in the two cases, it can be seen that an adjustment of the time scale can map synchronous and asynchronous applications to the same analysis method - and the results indicate that interaction can be detected in both cases;
- **Spatial versus workspace structure:** The application of a spatial interaction model in the puzzle game is natural, because the game is based on a spatial order of pieces. In the case of BSCW, this spatial structure does not exist, yet a simple mapping transforms the semantic structure of the workspaces to a spatial structure and makes it available for analysis. For the calculation of relevant spatial distances it is required to compute a baseline for each CSCL context used – as space varies depending on the semantic structure of the educational content and the organization of the artifacts (e.g., a significant difference in space is not the same in a drawing tool as compared to a file sharing system).
- **Success:** In the puzzle, one can calculate the success of the task, whereas this complicated in the BSCW environment since the inherent semantics of the artifacts is more complex. Nevertheless, a calculation of the relative *impact* – based on replies (i.e., whether a contribution evokes a response) – can be included in the mathematical model. The more interaction feedback an activity evokes, the higher is its assumed impact. In this case it is assumed that a higher degree of impact signals that a specific contribution is perceived as more influential by the group members (although this cannot objectively be assessed from log data). Viégas, Wattenberg and Kushal (2004) proposed methods for determining the impact of single user's contributions to co-authored documents. The same methods could be added to the calculations proposed in this paper to better asses the users' activities. Activities with larger impact would then lead to larger peaks in the IA values. Higher average *IA(t)* values do not automatically result in better learning. If interaction is intended in the didactical model, *IA(t)* provides a means to measure whether or not the students complied to this model. Best practices for learning groups can provide a baseline to work towards interpretation guidelines.
- **Group rhythm:** Since the duration of interaction was much longer for the BSCW groups, group interaction rhythms can be observed. In the puzzle game, such a rhythm is less significant although puzzle groups were observed that showed a rhythmic interaction. An analysis of several consecutive games might reveal group interaction for the puzzle game as well.

The examples showed that the interaction value revealed group rhythms in BSCW and the relation between interaction and success in a cooperative puzzle game. The calculation of the interaction value can assist theory building and evaluation to better understand group behavior. The interaction value metric appears especially relevant in those instances where a teacher simultaneously supervises multiple groups and time simply does not allow an in-depth supervision of each group. Although the calculation is based on log files, which are by nature limited in their explanatory power, the proposed interaction value metric extends current metrics that focus on effects *of* the system. It instead it is a metric that approximates interaction between users *with* the system.

It is a simplification to consider all artifacts as generic information objects and rate them all as equally important, as technically the metric could be applied to static educational websites where users are not aware of each other. Whether or not this simplification can be valid in systems that provide different kinds of information objects needs to be determined. Probably, it could make sense to analyze interaction differently according to the different topic areas of artifacts (e.g., having one analysis for calendar entries and another one for pages). It also can be argued that, for example negotiating a meeting (in a nice Irish pub) can 'mislead' the interaction value calculation. Yet, in a learning context social interaction is also important (cf. Schümmer & Haake, in press) and often it is taken for granted in CSCL (Kreijns, Kirschner, & Jochems, 2003). Future research includes the application of the model to discussion groups at the FernUniversität in Hagen to investigate highly interactive courses (in a CSCL-Setting) compared to low interactive courses (within a larger user community) and the difference between moderated and un-moderated discussions and the difference between discussions with long and short discussion threads. Another possible direction is to construct scenarios, such as collaboration patterns of individual members within a group that focuses on the same artifact at the same time, collaboration within subgroups, collaboration within a subgroup and individual work of some other users, and equality of participation in relation to semantically distributed artifacts (e.g., postings in different discussion forums). Finally, the metric could be extended to reveal that users interact differently with group members and a single users' impact on the interaction value, and thus constitute a representation of interactivity that can be generated on demand. Either the group can use this information to coordinate their collaboration or a teacher might use this information to guide specific intervention in a group (e.g., this representation enhances the overall interactivity on the level of group and provides the means to look more closely at the interaction to make a specific decision, for example contacting a less active – possibly free-riding – group member).

Despite the differences between the environments, it has been illustrated that the interaction value can be calculated. This paper has presented first applications of the measure *IA(t)* and the calculated curves revealed promising insights into group interaction. The interaction value can provide information that can be used by groups to coordinate their cooperation or by a teacher to supervise and/or asses the collaboration practice. At present the interpretation of the visualizations requires expert analysis. A tool to assist the application of the metric and automated generation of diagrams for *IA(t)* is currently being developed, which includes generating extreme cases for comparison to aid interpretation of the curves. Yet, further validation of the metric requires a systematic triangulation of log file data, IA value calculation and observations or interviews with students or focus groups. To prove its validity the measure needs to be applied to larger collections of log files (permitting statistical tests). The CSCL community can perform a leading role through applications of the measure.

Acknowledgements

Thanks are due to Dieter Becherer-Ecker , Jan Schümmer, Thorsten Holmer, Jörg Haake, Anja Haake, Alejandro Fernandez, and the OpenCOAST development team for helping with the puzzle experiments or providing valuable comments on earlier versions of this paper.

References

Begole, J. B., Tang, J. C., Smith, R. B., & Yankelovich, N. (2002). Work rhythms: analyzing visualizations of awareness histories of distributed groups. In *Proceedings of CSCW 2002* (pp. 334-343). New York: ACM Press.

Benbunan-Fich, R., & Hiltz, S. R. (1999). Impacts of asynchronous learning networks on individual and group problem solving: A field experiment. *Group Decision and Negotiation*, *8*, 409-426.

Bentley, R., Appelt, W., Busbach, U., Hinrichs, E., Kerr, D., Sikkel, K., Trevor, J., & G. Woetzel, G. (1997). Basic support for cooperative work on the world-wide web. *International Journal of Human-Computer Studies*, 46, 827-846.

Chan, C. K. K., & van Aalst, J. (2004). Learning, assessment and collaboration in computer-supported environments. In P. Dillenbourg (Series Ed.) & J. W. Strijbos, P. A. Kirschner & R. L. Martens (Vol. Eds.). *Computer-supported collaborative learning: Vol 3. What we know about CSCL: And implementing it in higher education* (pp. 87-112). Boston, MA: Kluwer Academic/Spinger Verlag.

Chyng, Y. J., Steinfeld, C., & Pfaff, B. (2000) Supporting awareness among virtual teams in a web-based collaborative system: The TeamSCOPE system. *ACM SIGGROUP bulletin*, *21*(3), 28-34.

Conklin, E. J., & Weil, W. (1997). *Wicked problems: Naming the pain in organizations*. Retrieved November 6, 2003, from http://www.touchstone.com/tr/wp/wicked.html.

De Graaff, R., De Laat, M., & Scheltinga, H. (2004). CSCL-ware in practice: Goals, tasks and constraints. In P. Dillenbourg (Series Ed.) & J. W. Strijbos, P. A. Kirschner & R. L. Martens (Vol. Eds.). *Computer-supported collaborative learning: Vol 3. What we know about CSCL: And implementing it in higher education* (pp. 201-219). Boston, MA: Kluwer Academic/Spinger Verlag.

Entwistle, N., & Tait, H. (1990). Approaches to learning, evaluations of teaching, and preferences for contrasting academic environments. *Higher Education*, *19*, 169-194.

Gentner, D. & Stevens, A. (1983). *Mental Models*. Hillsdale, NJ: Erlbaum.

Hannas, L. (1981). *The jigsaw book*. New York: The Dial Press.

Harasim L. (1993). Collaborating in cyberspace: Using computer conferences as a group learning environments. *Interactive Learning Environments*, 3, 119-130.

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

Ivory, M. Y., & Hearst, M. A. (2001). The state of the art in automating usability evaluation of user interfaces. *ACM Computer Surveys*, *33*, 470-516.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: A review of the research. *Computers in Human Behavior*, *19*, 335-353.

Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' online science discussions. *Learning & Instruction*, *13*, 487-509.

Pinelle, D., & Gutwin, C. (2000). A review of groupware evaluations. In *Proceedings of WET ICE 2000* (pp. 86-91). IEEE Computer Society.

Rodden, T. (1996). Populating the application: A model of awareness for cooperative applications. In *Proceedings of CSC 1996* (pp. 87-96). New York: ACM Press.

Rogers, C., & Freiberg, H. J. (1994). *Freedom to Learn* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Salomon, G., Perkins, D., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*, *20*, 9-20.

Schümmer, T. (2004). GAMA - A pattern language for computer supported dynamic collaboration. In K. Henney & D. Schütz (Eds.), *EuroPLoP 2003, Proceedings of the 8th European conference on pattern languages of programs* (pp. 53-113). Konstanz: UKV.

Schümmer, T. (2002). Enabling technologies for communities at web shops. In J. Plaice et al. (Eds.), *DCW 2002* (pp. 253-265). Heidelberg: Springer Verlag.

Schümmer, T., & Haake, J. (in press). Kooperative Übungen im Fernstudium: Erfahrungen mit dem Kommunikations- und Kreativitätswerkzeug FUB. In M. Beißwenger & A. Storrer (Eds.), *Chat-Kommunikation in Beruf, Bildung und Medien* [Chat communication in profession, education and media]. Stuttgart: ibidem.

Scouller, K. M. (1997). Students' perceptions of three assessment methods: Assignment essay, multiple choice question examination, short answer examination. *Research and Development in Higher Education*, *20*, 646-653.

Smith, A. (2000). Machine mapping of document collections: The Leximancer system. In *Proceedings of the fifth Australasian document computing symposium*. Sunshine Coast, Australia: DSTC.

Stahl, G. (2001). Rediscovering CSCL. In T. Koschmann, R. Hall & N. Miyake (Eds.), *CSCL 2: Carrying forward the conversation* (pp. 169-181). Mahwah, NJ: Lawrence Erlbaum Associates.

Stahl, G. (2004). Building collaborative knowing: Elements of a social theory of CSCL. In P. Dillenbourg (Series Ed.) & J. W. Strijbos, P. A. Kirschner & R. L. Martens (Vol. Eds.), *Computer-supported collaborative learning: Vol 3. What we know about CSCL: And implementing it in higher education* (pp. 53-85). Boston, MA: Kluwer Academic/Spinger Verlag.

Strijbos, J. W., Kirschner, P. A., & Martens, R. L. (2004). *What we know about CSCL: And implementing it in higher education*. Boston, MA: Kluwer Academic/Spinger Verlag.

Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (in press). Content analysis: What are they talking about? *Computers & Education*.

Viégas, F. B., Wattenberg, M., & Kushal, D. (2004). Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the 2004 conference on human factors in computing systems* (pp. 575-582). New York: ACM Press.