Time series analysis of collaborative activities

Irene-Angelica Chounta, Nikolaos Avouris

HCI Group, University of Patras, Greece
{houren,avouris}@upatras.gr

Abstract. Analysis of collaborative activities is a popular research area in CSCW and CSCL fields since it provides useful information for improving the quality and efficiency of collaborative activities. Prior research has focused on qualitative methods for evaluating collaboration while machine learning algorithms and logfile analysis have been proposed for post-assessment. In this paper we propose the use of time series analysis techniques in order to classify synchronous, collaborative learning activities. Time is an important aspect of collaboration, especially when it takes place synchronously, and can reveal the underlying group dynamics. Therefore time series analysis should be considered as an option when we wish to have a clear view of the process and final outcome of a collaborative activity. We argue that classification of collaborative activities based on time series will also reflect on their qualitative aspects. Collaborative sessions that share similar time series, will also share similar qualitative properties.

Keywords: time-series, collaboration, classification, logfile analysis

1 Introduction

The analysis of collaborative activities is a complex task due to the nature of collaboration itself and the amount of information that has to be evaluated. However in computer supported collaboration, all the information regarding users' interaction is recorded in logfiles by the groupware applications that mediate such activities. Therefore logfile analysis, automated metrics and other quantitative methods are used for post-assessment of the quality of collaboration, to trace any possible drawbacks and reveal underlying mechanisms that may affect the process and outcome of collaborative activities [1-3]. Most of these methodologies take into account the aspect of time. For example, how turn taking mechanisms affect communication, whether large gaps in the communication flow might be considered as a failure or large periods of individual work phases might affect coordination [4]. We argue that such phenomena can be captured using time series [5]. In this study we use sequences of events of collaborative activities to form time series that represent how the process unfolds in time, related with quantitative assessments of collaboration quality. We explore whether collaborative sessions that share similar time series characteristics are also of similar quality. To that end, predictions of quality of collaboration based on time series techniques were compared to assessments by evaluators for a rich dataset of collaborative activities. The use of time series as a tool of analysis of collaborative activities will add up and empower existing machine learning techniques while the workload of human evaluators will be minimized. Moreover, through time series techniques, real time assessment of the activity may be achieved. In that case, the evaluator will be aware, in real time, whether an activity is turning out successfully or, otherwise, which collaborative aspect should be further supported.

This paper is organized as follows. In section 2 the time series construction from the logfiles of various collaborative activities is described. In section 3 we discuss the structure and techniques used for the memory-based classification model that is proposed. The construction of the model itself is also analyzed. The results of the study are presented in section 4 and in section 5 we conclude with a general discussion about the setup and results of this study as well as improvements and future work.

2 Collaborative activities as described by time series

A computer-supported collaborative activity can be described via a multivariate time series of events (such as chat messages). We propose the use of multivariate time series because the way an activity builds up in time and its cross-correlation with other activity that occurs concurrently are important. To fully explore the underlying mechanisms and dynamics of collaboration such information might be proven useful and should not be ignored.

Time series is defined as any sequence of observations recorded at successive time intervals. Network traffic monitored by a web server per hour or the price of shares in a stock market per week are examples of time series commonly used and analyzed for various purposes. Time series fall into two categories: univariate and multivariate. A multivariate time series is a vector of more than one time series which are cross-related. The objective of time series analysis is to gain understanding of the nature and underlying mechanisms of a monitored activity, to group and classify samples based on their time series properties and to forecast. Many models and algorithms have been proposed to deal with univariate or multivariate time series analysis and classification, such as ARIMA, VAR, Hidden Markov Models (HMM), Dynamic Time Warping (DTW) and Recurrent Neural Networks. Time series analysis is widely used in a variety of fields such as economics, biology and computer science [6], [7].

For the purposes of this study we used the logfiles of collaborative activities that took place during a programming course in the Dept. of Electrical and Computer Engineering. The subject of the course was the joined construction of flow charts by dyads. The groupware application that mediated the activities provides users of a common workspace for the construction of diagrammatic representations and a chat tool to support communication between partners [8]. All activity was recorded in logfiles for later use. The setting of the study was such that same conditions applied for all clients/collaborators (e.g. equal numbered groups, high speed local area network, identical computers) and similar network delays occurred for all clients.

In order to portray a collaborative session as a time series we computed the sum of chat and workspace activity of collaborating partners per time intervals and per events. The sequences of aggregated chat and workspace events form up a multivariate time series per collaborative session. Previous studies show that a number of metrics regarding chat and workspace activity highly correlate with the quality of collaboration of a joined activity [9]. Based on these studies we made use of the metrics displayed in Table 1 to assemble time series for each session. Therefore each collaborative session is represented by one multivariate time series which is practically a vector of four univariate time series constructed from aggregated chat and workspace events, where:

- Number of chat/workspace actions per time interval: the sum of messages/workspace actions of both partners in a time interval.
- Roles' alternations in chat activity per time interval: the number of times the active role of a partner was switched in chat/workspace activity in a time interval.

Table 1. Chat and Workspace metrics used for time series construction

Metric	Sums	Difference of Sums	Alternations	Difference of Alternations
chat	number of chat messages per time interval	difference of chat messages between consecutive time intervals	roles' alternations in chat activity per time interval	difference of roles' alternations in chat activity between consecutive time intervals
workspace	number of workspace actions per time interval	difference of workspace events between consecutive time intervals	roles' alternations in workspace activity per time interval	difference of roles' alternations in workspace activity between consecutive time intervals

Another critical point when creating time series from aggregated events is the time interval chosen [10]. Valuable information might be lost in case of a small or large time interval. The choice of the appropriate time interval is a critical task and domain dependent. Therefore a variety of time intervals was tested before concluding. In this study, the duration of collaborative activities ranged between 50 minutes to 1 hour and a half. The time intervals studied were of 1, 5, 8 and 10 minutes. Activity for time intervals of less than one minute was not explored since the number of events occurring within such time periods was small.

3 Methodology of Analysis

The aim of this study is the classification of collaborative sessions using their time series properties. For this purpose we create a data pool of time series, extracted from collaborative activities, associated with quantitative assessments of collaboration quality. The suggested set up is fashioned after memory-based learning models using time series [11]. Memory-based learning presupposes the use of a distance measure. For that we used the Dynamic Time Warping (DTW) algorithm.

3.1 Dynamic Time Warping

DTW is an algorithm that measures similarity between two sequences that vary in time. It provides a distance measure termed DTW distance. Originally it was used for sound and video processing but has also found many applications in time series analysis [12], [13]. We used the DTW algorithm implementation proposed by Giorgino, T., for the R-statistics software [14]. This implementation allows choosing between multiple options for step patterns (the way consecutive time series elements are matched) and dissimilarity functions (for the cross-distance matrix computation). The DTW algorithm does not presuppose time series' stationarity or non-missing information as Fourier transform does and this is one of the reasons for its growing popularity.

3.2 Quality of collaboration

The rating scheme proposed by Kahrimanis et al. [3] was used to provide quantitative judgments of the quality of collaboration for the sessions used in this study. The aforementioned rating scheme proposed the rating of seven collaborative dimensions on a 5 point scale, that stand for the five, fundamental aspects of collaboration: communication, joint information processing, coordination, interpersonal relationship and motivation [4]. These seven dimensions are: collaboration flow, sustaining mutual understanding, knowledge exchange, argumentation, structuring the problem solving process, cooperative orientation and individual task orientation. The rating was carried out by two trained evaluators. We made use of the average value of six out of seven, dimensions leaving out the motivational/Individual task orientation aspect which is rated for each student separately. We denote this metric as Collaboration Quality Average (CQA) and it takes values within the range {-2,2}. As stated in Section 2, CQA has been found highly correlated with logfile metrics of interaction. Therefore we argue that similar time series will have similar CQA evaluative values.

3.3 Memory-based classification model

The memory model construction and classification procedure consists of three steps.

1. Time series construction from the logfiles of collaborative sessions. For each collaborative session we constructed its multivariate time series representation, as described in section 2. Outliers were detected by visual inspection of time plots and deleted from the final dataset.

2. Input of sample entries in memory. The data pool consists of the multivariate time series extracted from collaborative activities, 212 samples in total, as collected in step 1. Each sample's quality of collaboration is also assessed by two evaluators (section 3.2). Therefore each point in the memory stands for a collaborative session and is described by its time series and an evaluation value for the quality of collaboration (CQA).
3. Classification of a query sample. By the term "query sample" we name any collaborative session that is not accompanied by an evaluation value CQA. The purpose is to approximately estimate the evaluation value CQA by finding the optimal time series match of the session among the sessions of the data pool. We argue that collaborative activities described by similar time series will have a similar CQA evaluation value. Therefore if the time series of two samples ts_Sa and ts_Sb, with Sa being the query sample and Sb the reference sample, have a minimum distance DTW then the evaluative value CQA of Sa should be approximately equal to the evaluative value CQA of Sb.

4 Results

The dataset used initially consisted of 228 collaborative sessions. The logfiles of the sessions, as recorded by the groupware application that mediated the activity, were used for the construction of the multivariate times series of aggregated events. Outliers were removed and the final dataset used in the memory-based classification model consisted of 212 collaborative sessions. For each one of the 212 samples we computed the DTW distance and found the optimal match from the 211 samples remaining in the data pool. The study was repeated for a variety of time intervals in aggregated events (1, 5, 8 and 10 minutes), two dissimilarity functions (Euclidean and Manhattan) and two step patterns (symmetric1 and symmetric2) of the DTW algorithm. In order to evaluate the results, as well as define the most appropriate time interval, dissimilarity function and step pattern, we estimated the correlation matrix of the evaluative value CQA (predicted vs. true value), the root mean squared error (RMSE) and the mean absolute error (MAE) for each case.

Correlation is a popular method to explore statistical relations between variables. Spearman's rank correlation coefficient is used to reveal any existing relations between the evaluative value CQA of a session, as assessed in the evaluation phase, and the predicted - by the memory-based classification model - value. For step pattern set to the symmetric $P = 0$, the two variables are significantly correlated for most of the combinations of time intervals and dissimilarity methods. Spearman's rho depicts the degree of the relation between two variables and it may range from -1 to 1. A value of 1 shows a strong, positive correlation while a value of -1 reveals a strong, negative. In our case the real and predicted evaluative values of CQA are positively and significantly correlated for all time intervals (Table 2). However the strongest correlation occurs for 1 minute time interval and Manhattan as a dissimilarity method ($p<0.05$, rho=0.3).

Table 2. Spearman's Rho correlation coefficient for CQA real and predicted values per time interval, for Manhattan and Euclidean dissimilarity methods and for step pattern symmetric P=0

	Manhattan		Euclidean	
time interval	p value	Rho	p value	Rho
1 minute	0,000	0,3	0,029	0,15
5 minutes	0,002	0,2	0,021	0,15
8 minutes	0,000	0,23	0,005	0,18
10 minutes	0,011	0,17	0,010	0,17

The mean absolute error (MAE) and root mean squared error (RMSE) are used to measure accuracy; how close the predicted values are to the "observed", or real, ones. For MAE individual differences are weighted equally in the average while large errors are highly weighted in RMSE. The difference of MAE and RMSE can be used as an insight of the variance of individual errors in the sample. We provide the results of analysis for both error metrics in Table 3. Smallest MAE and RMSE values occurred for 1 minute time interval and Manhattan method (MAE=0.89, RMSE=1.1) while largest values occurred for 5 minutes interval and Euclidean distance (MAE=1.21 RMSE=1.5). The variance of individual errors is also minimized for 1 minute time interval. The extent to which the alignment of DTW avoids mismatches is influenced by the choice of the dissimilarity function (Euclidean or Manhattan) [14]. Therefore the difference portrayed in the results can be justified.

Table 3. MAE and RMSE for CQA real and predicted values per time interval, for Manhattan and Euclidean dissimilarity methods and for step pattern symmetric P=0

	MAE		RMSE		RMSE - MAE	
Time Interval	Manhattan	Euclidean	Manhattan	Euclidean	Manhattan	Euclidean
1min	0.89	0.97	1.14	1.21	0.25	0.24
5min	1.19	1.21	1.48	1.50	0.29	0.29
8min	1.18	1.16	1.50	1.48	0.32	0.32
10min	1.17	1.19	1.44	1.47	0.27	0.28

Both correlation and error measures reveal that time series of aggregated events per time interval of 1 minute portray better the quality of collaboration of a group activity as this has been assessed by human evaluators using a rating scheme. In case we use the same metrics for larger time intervals of the activity, valuable information might be lost (Fig. 1). For time interval of 1 minute, the distribution of absolute difference among the CQA predicted and CQA real value is shown in Fig. 2. The absolute difference for 41% of cases was less than 0.5, for 68.4% of the cases difference was less than 1 while the 92% falls below a difference of 2 points.

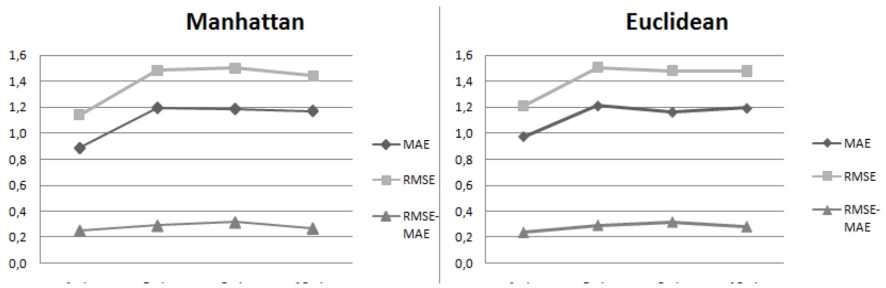

Fig. 1. MAE and RMSE per dissimilarity function

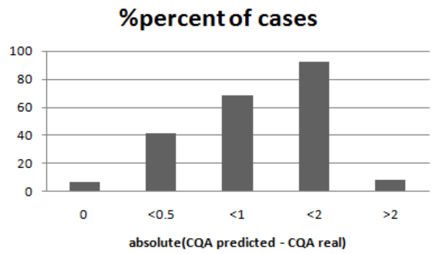

Fig. 2. Distribution of absolute difference among the CQA predicted and CQA real values

5 Conclusions and future work

In this paper we discuss the use of time series for classifying collaborative activities. The main goal of the study was to explore whether this classification regarding the time series characteristics of collaborative session depict also their qualitative assessments of collaboration quality. For the classification a memory-based approach was proposed and the Dynamic Time Warping algorithm was used as a dissimilarity measure between time series. The results revealed that there is a significant positive correlation among the predicted and real evaluative values (CQA) while the lower values for MAE and RMSE values as well as for the variance of individual errors in the sample occur for 1 minute time interval (0.89, 1.1 and 0.25 respectively, for a value range {-2, 2}). The error metrics MAE and RMSE are used to provide insights of a model's accuracy. In order to have a more complete picture we should also take into account qualitative aspects such as the simplicity of the model and the context within it is used.

In the classification procedure one optimal match was used for each query time series. Results could be improved if we used more advanced techniques such as k-nearest neighbor or applied weights in the DTW algorithm. Moreover other classification models for multivariate time series should be studied, such as Hidden Markov Models [15], as well as popular clustering methods like hierarchical clustering and kmeans. In future work, a qualitative analysis regarding time series characteristics and how they are related with individual aspects of collaboration will be explored, as well as how random shocks, such as communication failure due to network breakdowns, affect the collaborative process and the way partners recover from that.

6 References

1. Barros, B., Verdejo, M. F.: Analysing student interaction processes in order to improve collaboration. The DEGREE approach. In: International Journal of Artificial Intelligence in Education, 11(3), 221-241 (2000)
2. Schümmer, T., Strijbos, J.-W., Berkel, T.: A new direction for log file analysis in CSCL: Experiences with a spatio-temporal metric. In: Koschmann, T., Suthers, D., Chan, T.W. (eds.) Computer Supported Collaborative Learning 2005: The next 10 years!, pp. 567–576 (2005)
3. Kahrimanis, G., Meier, A., Chounta, I.-A., Voyiatzaki, E., Spada, H., Rummel, N., Avounis, N.: 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. 4th European Conferenceon Technology Enhanced Learning, EC-TEL 2009. Nice, France, September 29–October 2, 2009, Springer-Verlag, pp. 267–272 (2009)
4. Meier, A., Spada, H., Rummel, N.: A rating scheme for assessing the quality of computer-supported collaboration processes. In: International Journal of Computer-Supported Collaborative Learning, 2(1), 63–86 (2007)
5. Vasileiadou, E.: Stabilisation operationalised: Using time series analysis to understand the dynamics of research collaboration. In: Journal of Informetrics, 3(1), 36-48 (2009)
6. Box, G. E. P., Jenkins, G. M, Reinsel, G. C.: Time series analysis, Forecasting and Control, Fourth Edition, Wiley (2008)
7. Chatfield, C.: The Analysis of Time Series: An Introduction, Chapman & Hall (2003)
8. Avouris N., Margaritis M., Komis V.: Modelling interaction during small-group synchronous problem-solving activities: The Synergo approach. In: 2nd Int. Workshop on Designing Computational Models of Collaborative Learning Interaction, ITS2004, Maceio (2004)
9. Kahrimanis, G., Chounta, I-A., Avouris, N.: Study of correlations between logfile-based metrics of interaction and the quality of synchronous collaboration. In: International Reports on Socio-Informatics (IRSI), vol. 7(1), pp. 24–31 (2010)
10. Shellman, S. M.: Time series intervals and statistical inference: The effects of temporal aggregation on event data analysis. In: Political Analysis 12(1), 97-104 (2004)
11. Deng, K., Moore, A., Nechyba, M.: Learning to recognize time series: Combining arma models with memory-based learning. In: IEEE CIRA, pp. 246–250 (1997)
12. Warren Liao, T.: Clustering of time series data—a survey. In: Pattern Recognition 38(11), 1857-1874 (2005)
13. Chao, S., Wong, F., Lam, H-L., Vai, M-I.: Blind biosignal classification framework based on DTW algorithm. In: Machine Learning and Cybernetics (ICMLC) International Conference, vol.4, pp.1684-1689 (2011)
14. Giorgino, T.: Computing and visualizing dynamic time warping alignments in R: the dtw package. In: Journal of Statistical Software 31(7), 1–24, (2009)
15. Kirshner, S.: Modeling of multivariate time series using hidden Markov models. PhD thesis, University of California, (2005)