

Using Sequence Analysis to Characterize the Efficiency of Small Groups in Large Online Courses

Dorian DOBERSTEIN ^{a*}, Tobias HECKING^a & H. Ulrich HOPPE^a

^a*COLLIDE Research Group, University of Duisburg-Essen,*

Lotharstr. 63, 47057 Duisburg, Germany

*doberstein@collide.info

Abstract: Small group collaboration can enrich the learning experience in large online courses in the interactional and social dimensions. The experience underlying this study stems from an inter-university online course with collaborative writing as a group task. Here, coordination between group members was supported through a discussion forum and collaborative writing was facilitated through a shared web-based editor. Actions in the forum and in the writing tool were categorized as coordination, monitoring, minor/major contribution. Pair-wise similarities between the corresponding action sequences were captured in a similarity matrix, which formed the basis for a cluster analysis. The clusters show specific patterns especially regarding the distribution of inactivity and coordination. Our findings show that inactivity can be counterbalanced by early coordination without rendering a group dysfunctional.

Keywords: analysis, learning groups, online courses, collaboration patterns

1. Introduction

In MOOCs and other types of online learning courses, the basic activities are video watching and the completion of assignments such as self-test quizzes. This is consistent with the rationale to support individual learners in self-directed knowledge acquisition independent of location and time. In these scenarios, collaboration typically plays a minor role and is often limited to the interaction in discussion forums. A typical problem of those courses is a lack of individual support and thus a lack of incentives as they might arise from social interaction in a shared environment.

It has been argued that online courses should be further adapted to individual learner needs by offering learning activities and assistance taking into account specific needs or profiles of learners or learner types (Grünewald, Meinel, Totschnig, Willems, 2013). Given a high number of participants and limited resources on the side of teaching staff, practical solutions rely on peer-to-peer interaction to facilitate collaboration and group work (Wichmann et al., 2016; Ferschke et al., 2015; Staubitz et al., 2015). This can also help to establish an effective learning community in which the lack of individual support is compensated by decentralized peer-help and self-organized discussions. Recent research provides evidence that collaboration and a sense of being part of an active community can also reduce attrition in online courses (Tomar, Sankaranarayanan, Rosé, 2016). Furthermore, the heterogeneity of background knowledge and points of view in a large audience can be exploited for different kinds of group compositions and to facilitate knowledge exchange and critical discourse between participants (Wichmann et al., 2016).

Recently, the effect of different strategies and algorithms for composing small learning group based on individual learner profiles or student models has gained considerable attention (see, e.g., Konert, Burlak, Steinmetz, 2014). However, from a practical perspective, the issue of establishing well-functioning and productive learning groups is at least of equal importance. Consequently, strategies for group formation should be combined with strategies for supporting group work, and the latter may even be more crucial. In pure online courses, group work often takes place asynchronously and communication is mediated and constrained by available technologies.

Typical problems of group work such as social loafing and a lack of commitment of the members can become more crucial in online courses due to anonymity and limitations of communication facilities (Piezon, Donaldson, 2005). This is very problematic since low productivity or even inactivity of single members negatively affects the learning experience of the other members. Longer periods of inactivity can cause uncertainty about the willingness of group members to participate. Limitations of social presence in online courses further complicate this issue (Roberts, Lowry, Sweeney, 2006; Weinel et al., 2011).

Especially in the context of CSCL, the analysis of temporal aspects in learning activity sequencing has already gained some interest (Reimann, 2009). Our target is a fine-grained temporal analysis and characterization of different types of learning groups in an online course that featured participation in group discussions and collaborative writing as core activities. The goal is to identify characteristic patterns in those collaborations. On the input level, our approach relies on human coding to categorize actions in the form of different contribution types: coordination, monitoring, minor contribution, and major contribution. In addition, one-day inactivity is introduced as another descriptor. Sequences of such action descriptors are then collected for portions of group work corresponding to a specific task assignment. A similar approach for analyzing user activity sequences in online courses has recently been presented by Shirvani Boroujeni and Dillenbourg (2018) who link activity sequences of users to their overall studying behavior. The basic activities in this approach are related to video watching and the submission of assignments throughout the course of online lectures. A distinction is made between a hypothesis driven approach, which tries to detect predefined activity patterns that correspond to certain learning styles, and a data driven approach using unsupervised learning methods to facilitate an open-ended detection of any kind of procedural learning patterns. In the first approach, one activity pattern (or sequence) corresponds to one behavioral state. Transitions between such states represent changes in the strategy over time (between assignments). Our work utilizes unsupervised learning methods without predefined activity patterns, similar to Shirvani Boroujeni and Dillenbourg's data driven approach. The basic elements here are single actions and not predefined action sequences.

The specificity and novelty of our approach lies in the usage of sequence analysis techniques that originate from bioinformatics. Originally, these methods have been used to determine similarities or differences in DNA, RNA or peptide sequences based on an "edit distances" (i.e., the cost of transforming one sequence into another). Meanwhile they have been applied to a variety of applications, including social interactions (Cornwell, 2015). In our context, the pair-wise similarities (based on edit distances) between the action sequences are captured in a similarity matrix that forms the basis for a cluster analysis. The clusters show specific patterns especially regarding the distribution of inactivity and coordination, and they can also be compared in terms of group productivity, quality of the results and work distribution.

The results of the analysis described in this paper are supposed to contribute to a better understanding of group work in online courses and to highlight possible starting points for the development of proper intervention mechanisms and support mechanisms based on early identification of collaboration problems.

The remainder of this paper is organized as follows: Section 2 describes the background of this study in more detail and relates it to existing research. Section 3 outlines the approach for our analysis, Section 4 presents the results and the last section discusses the findings and future work.

2. Background: Group Work in Large Online Courses

In order to test different approaches of introducing group work in large online courses two subsequent instances of a lecture on "computer mediated communication" were conducted and analyzed (Course 1: N=270; Course 2: N = 111). These courses were open to students of different study programs from two universities and participants could receive credits. The design of the courses was inspired by contemporary MOOCs and the course language was German for both courses. The course platform based on Moodle was adapted to the needs of collaborative online courses. In each course section, students were provided with a short instructional video, which covered the most important aspects of the topic, literature and self-test quizzes to acquire theme specific knowledge. In addition, the students received assignments that had to be accomplished in

small groups of four participants each during one week. The goal of these assignments was to collaboratively create a short text for a given topic. Two existing Moodle plugins were adapted to support the groups in solving the tasks. For text creation students used the collaborative editor Etherpad, integrated in the learning platform and thus enabling real-time collaboration. Discussions and coordination activities were supported by separated discussion forums for each group. Discussion forums and Etherpads were linked such that the students could constantly switch between the two.

One of the experiences of the first course was that it is difficult to maintain an adequate level of activity of the learning groups. Activity gaps (longer inactive periods) in the group work were identified as a serious problem. In worst cases particular group members felt uncertain about the group still being “there” and active, even if the list of group members was visible all the time. Furthermore, a relationship between the satisfaction with group work and overall satisfaction with the course could be found (Kyewski et al., 2016). Based on these first experiences, significant improvements in course satisfaction could be achieved in the second instance of the course by a clearer structuring of the group activities and more strict guidelines for solving group tasks (Erdmann et al., 2017).

3. Analysis

In this analysis, a dataset was used that contains the activities of 81 students which participated in the group tasks in the second course mentioned in Section 2. The tasks were to write 600-word wiki articles on different topics. There were five consecutive group tasks with a duration of one week each. For each task new groups containing 4 students each were formed. A similar analysis approach to the one presented here was first applied to a smaller dataset containing the activities of 19 groups resulting from a strategic group formation (Doberstein, Hecking, Hoppe, 2017) The basic set of groups was now extended to contain the 65 groups that were mainly composed at random. In addition to enlarging the dataset, the evaluation was extended using content analysis on the results of the groupwork and an analysis of the distribution of text contributions among the group members.

The first and most time-consuming step in our analysis was data pre-processing. For each learning group, the traces of individual contributions in the discussion forum and Etherpad were assembled into action sequences based on their temporal order. Furthermore, each contribution in these sequences was manually classified according to the nature of the contribution. This resulted in one encoded collaboration sequence per group. In a second step, similarities between these sequences of encoded actions were calculated. Based on these similarities, a cluster analysis was performed. Afterwards, the relation between the cluster affiliation, the productivity of the group, the quality of the wiki articles and the work distribution between group members was measured.

For the pre-processing, several steps were necessary. The first step selected the contributions in the group forum and the text written in the Etherpad editor which were logged in the database for each group and ordered them temporally. Since the Etherpad is a real-time editor without information on revisions, all characters written by a single user subsequently without a break of more than 60 minutes were subsumed as a single contribution. The results were action sequences of forum posts and text snippets contributed in the collaborative editor for each group in chronological order.

Next, the single actions of each sequence were classified into four main categories (see Table 1). Contributions in the Etherpad were classified as either major contributions or minor contributions. Major contributions added a considerable amount of text (>600 characters) and extended the semantic content of the text. Minor contributions were small improvements in spelling or smaller text modifications (<600 characters). For the most part, posts in the forum were classified as either coordination or monitoring. Posts were part of the coordination class if they were dedicated to organizing the groupwork. These were messages with a prospective character, for example planning or work distribution. On the other hand, retrospective posts were classified as monitoring. These were for example reports regarding own contributions to the Etherpad, the status of the groupwork or technical problems. Though the forum was meant for communication between the group members while the Etherpad was meant for topic related activities, occasional students would post their contributions to the forum to discuss them. In that case forum posts could also be classified

as major or minor contribution according to the rules mentioned above. This classification of the actions was done manually which was time consuming. Occasionally actions occurred that were not related to the group task. These other actions were deleted from the data.

The resulting encoded collaboration sequences only contained actions of the group members but did not reflect phases of inactivity during which no group member showed any activity in the forum or the Etherpad. These passive phases were of special interest since they have a negative impact on the group work (see Section 2). In order to include these phases of inactivity a new gap action was inserted into the collaboration sequences whenever there was an inactive period of 24h. To identify inactive periods before the first activity of a group member took place, a start element was added as first item in each sequence with a timestamp that corresponds to the start time of the assignment.

Table 1

Coding Scheme for Contributions

Contribution type	Description	Examples
Coordination	Forum posts of prospective character, e.g. planning, distribution of work, commitments for envisaged contributions according to own time schedule.	“I could write something in the introduction, but I won’t have the time until tomorrow.” “Do you have some ideas how to structure the text?”
Monitoring	Retrospective forum posts, e.g. reports of contributions, reflection on the progress.	“I wrote something in the discussion part. Could you have a look if it fits?”
Major Contribution	Contributions in Etherpad with more than 600 characters that significantly expand the content of the text. Posts in the group forum that are supposed to be integrated into the Etherpad text.	
Minor Contribution	Small changes such as restructuring of the text, correction of typos, and other small text additions below 600 characters, or drafting an outline.	
Other	Forum posts that cannot be considered as relevant for the product of group work.	“The text looks good to me, I have nothing more to contribute”. “Thank you for the good work”.



Figure 1. Complete collaboration sequence for one group.

The resulting encoded collaboration sequences of the groups could then be interpreted as characteristic fingerprint of the collaboration activities of each group during an assignment. Figure 1 shows such a sequence for one group. Each action is colored according to its class. The sequence begins with the start element (red). Subsequently two gaps follow. After that all different contribution types appear in the sequence, which ends with a minor contribution.

In the next step, the final group sequences were compared, to find similarities between them. Furthermore, four properties regarding the quantity and quality of the resulting wiki articles were calculated.

3.1 Sequence matching and clustering

The sequence matching approach begins with the calculation of a distance measure followed by similarity-based clustering on the basis of the ensuing distance matrix. A proper measure for the distance of two collaboration sequences can be derived from sequence alignment or optimal matching as used in bioinformatics. Abbott and Tsay (2000) review the application of similar techniques in social studies. The idea of optimal matching is to calculate the minimal costs of transforming one sequence into another by insertion, deletion and substitution of sequence elements (similar to the Levenshtein distance). Here, different operations can be loaded with different basic costs. For our analyses of collaboration sequences, we defined the cost for insertion and deletion as 1 while the cost for substitution was 2 if a gap was involved and 1 in all other cases. The reason is to emphasize the difference between inactivity and active contributions of users. For example, changing a minor contribution into a major contribution to match one sequence to another does not make a big difference considering the course of actions in the development of a shared document. However, changing a gap into an action or an action into a gap should be more expensive since inactive periods can be an indicator for problems in the group work and as such substituting them should be weighted accordingly.

The resulting distance matrix was used to group the sequences into clusters. For the clustering the method Partitioning around medoids (PAM) was used (Kaufman, Rousseeuw, 1987). The idea of PAM is to search for k representative instances (medoids) as representative for each cluster. The algorithm first searches for a suitable set of representative instances (build phase) and then matches the other instances to their closest representative until no switch of objects between clusters improves the results. The clustering was performed with different number of clusters k , ranging from 2 to 4. To determine which clustering was best, the measures of diameter of the clusters, average distance inside the clusters and average distance between the clusters were taken into account. $K=3$ performed best in all categories.

3.2 Wordcount, work imbalance, and concept coverage

To rate the result of the produced wiki articles four properties were examined. The first property, the 'wordcount' is based solely on the productivity of the group and simply measures the number of words in the final version of the produced text. It is important to note that in all group assignments the students were supposed to write at least 600 words. Thus, taking the wordcount only for assessing the outcome produced by a group is of limited significance. However, some groups exceeded this minimum by hundreds of words. Accordingly, the wordcount can still be used to differentiate between groups that solved the task and groups that were more engaged.

To better judge the actual content of the produced texts over sheer wordcount, the wiki articles were compared to transcripts of the instructional videos that were provided for each thematic section of the course (see Section 2). Since these instructional videos covered the most important aspects in a course section, texts produced by good groups were assumed to be more similar to the video transcripts than others. For text comparison, part-of-speech tagging and word stemming was used to reduce both the wiki articles and the video transcripts to the contained nouns. More frequent nouns in the video transcripts can be considered as core concepts. The frequency vectors of the remaining terms of the articles and the video transcript of the corresponding course section were compared using cosine similarity. This similarity is called 'concept coverage' in the following. In order to make the computed cosine similarities between different course sections (topic of the week) comparable (topics differ in the number and distribution of core concepts), a ranking of groups was generated for each week. The group with the highest concept coverage was ranked 1, the one with the lowest was assigned the lowest rank. Since the range of the ranks depended on the number of articles that were written on the different topics (e.g. 19 articles on brainstorming, 9 articles on common ground), the ranks were normalized to range between 0 and 1. 0 means the group had the worst rank for the task assignment and 1 being the highest rank.

The third and fourth property subsumed as ‘work imbalance’ aim at how balanced the distribution of text contributions is among the group members. First, the share of characters each group members has contributed to the overall text was calculated. The variance between these shares within the groups gives an insight on how fair the work was distributed. A high variance means in this case, that the contribution to the text was unevenly distributed while a low variance indicates a well distributed group work. For this calculation, only students that were active and contributed to the task were included.

Alternatively, the Gini-index of the contributed characters of the 4 group members was calculated as well. The Gini-index ranges between 0 and 1 and measures the deviation of a distribution from an (ideal) uniform distribution. It is 0 if all group members contributed equally and 1 if one did all the work alone. Here also inactive group members were taken into account.

3.3 Decision tree induction

In order to further interpret the clusters and to investigate the possible effect of activity gaps and coordination messages, a decision tree has been constructed using the CART algorithm (Breiman, Friedman, Olshen, Stone, 1984). Here, each sequence was described by the number of the coordination messages and gaps while the corresponding cluster (cluster 1, 2 or 3) was used as dependent variable. The algorithm aims at building a compact tree, which best explains the association of sequence instances to clusters, by recursively splitting the instances into subsets according to the most discriminating property. Consequently, the more discriminating properties occur at higher levels of the tree. The procedure terminates when subsets become smaller than 5 instances. Later, post-pruning based on 10-fold cross validation (c.f. Breimann et al., 1984) is applied to reduce the tree to the most expressive split points.

4. Results

The clustering divided the 65 sequences into 3 clusters (see Figure 2). For each cluster the average number of produced words in the wiki article (Word count), the average number of phases of 24h inactivity (#Gaps), the average number of coordination activities (#Coordination), the concept coverage and the two versions of work imbalance are given in Table 2. Additionally, Table 3 shows the percentage of gaps and coordination messages in the first half of the group activity sequences. Generally speaking, the majority of the inactive phases can be found in the first half of the week for all clusters (51% - 86%), indicating that the groups need time to organize at the beginning and tend to work more closely towards the assignment deadline.

Table 2

Average Values for Produced Text, Number of Gaps, Number of Coordination Messages, Concept Coverage and Work Imbalance for each Cluster (Best Values are in bold face)

Cluster	Word count	#Gaps	#Coordination	Concept coverage	Work imbalance (variance)	Work imbalance (Gini index)
1	830	2.7	6.7	0.52	0.03	0.36
2	695	3.6	2.4	0.53	0.05	0.4
3	628	4.8	1.2	0.47	0.12	0.5

The 19 sequences in the first cluster showed the highest productivity with 830 written words on average in the wiki article. The majority of the groups start the work on the first or second day. Notable is the high number of coordination messages in this cluster. On average, each sequence contains 6.7 coordination messages with 68% of them in the first half of the assignment and only 2.7 activity gaps per sequence. The normalized average rank for the concept coverage (see Section 3.2) for cluster 1 and cluster 2 is very similar (0.52 and 0.53). Consequently, the groups in both clusters

cover the most important topics mentioned in the introduction video equally well while cluster 3 is slightly worse.

The work imbalance shows a similar picture but discriminates more between cluster 3 and the other two. The average variance of the text distribution for cluster 1 and cluster 2 (0.03 and 0.05) is below the average for all groups (0.06). This means that the contributions to the wiki article were more evenly distributed between the group members in clusters 1 and 2 than in cluster 3 (0.12). Over all groups, the variance ranged from 0.0, for a group where all members contributed equally to 0.44 where one student was responsible for most of the work. The work imbalance measured by the Gini-index gives a similar result. To further assess how strong sequence patterns indicate the distributions of work in the learning groups an ANOVA was conducted. Both measures for work imbalance (contribution variance and Gini-index) show a significant difference between the clusters ($F(2,62)=4.43$, $p=.02$) and ($F(2,62)=4.36$, $p=.02$) respectively. Following pairwise t-tests further revealed no difference between clusters 1 and 2 but significant differences between 3 and the other two with respect to the Gini-index and the variance-based measure.

Table 3

Percentage of gaps and coordination messages in the first half of the sequences for each cluster

Cluster	Gaps (first half)	Coordination (first half)
1	70%	68%
2	51%	47%
3	86%	19%

Cluster 2 consists of 17 groups that wrote 135 words less on average than cluster 1 groups. Groups in cluster 2 have 3.6 gaps on average, which is 0.9 more than the groups in cluster 1. A greater difference can be seen in the number of coordination messages which is 2.4 for the 2nd cluster. 47% of the coordination takes place in the first half of the activity sequence.

The biggest cluster is the 3rd one which consists of 29 sequences. These groups were least productive (628 words), showed the biggest number of inactive phases (4.8) and had the least coordination of all clusters (1.2). Nearly 2/3 of the groups in this cluster showed no activity in the first three days and especially coordination was rare in the beginning (only 19% of the coordination took place in the first half of the sequences).

Cluster 3 has the lowest average rank for concept coverage of all clusters (0.47). Furthermore, the groups in cluster 3 had the highest variance in text shares between the students (0.12). This means that the contributions to the wiki article were not as fairly distributed between the group members as they were in clusters 1 and 2.

Generally, it is expected, that groups with more inactive phases are less productive. However, the difference in the number of inactive phases between clusters 1 and 2 is just 0.9 while the word count of cluster 1 with an additional 135 words on average is much higher. One reason could be that the number of coordination messages is responsible for the disparity and that the impact of coordination especially in the beginning of the task is important for the productivity. While the wordcount for cluster 1 is higher, clusters 1 and 2 are similar in both concept coverage and text distribution.

As described in Section 3.3, a decision tree explaining the clusters based on the number of coordination messages and gaps was constructed (see Figure 3). Each non-leaf node denotes a split point that divides the sequences into two subsets. Every node shows the cluster that the majority of the sequences in the subset belong to at the top (also indicated by node color). Below the number for the cluster, each node contains three values that describe the fraction of instances in the split belonging to each of the three clusters (.00 = 0%; 1.00 = 100%; cluster 1-3 from left to right). The percentage at the bottom of each node shows how many of the entire set of sequences belong to the corresponding split. The higher a split point is in the hierarchy the more discriminating, and thus more important, is the split point. It can be seen that cluster 1 can almost completely be explained only by a number of coordination messages greater than 3.5 (78% of the instances). For those

sequences, that show less than 3.5 coordination messages, a second split point decides whether the groups are more likely to be in cluster 2 or cluster 3. This second split point divides the groups based on the number of gaps that appear in the sequence (more or less than 4.5). Those sequences with less than 4.5 gaps are more likely to belong to cluster 2 (56%) while the other sequences with more than 4.5 gaps most likely belong to cluster 3 (83%) A high number of gaps indicate the low productivity, the lowest concept coverage and the biggest work imbalance of cluster 3.

This observation highlights the importance of coordination messages. Groups that are well coordinated belong to the best performing cluster 1 while they cannot be well distinguished from cluster 2 based on the number of gaps. Thus, the presence of activity gaps does not necessarily go along with low productivity if enough coordination takes place. However, the absence of coordination seems to be coupled to the number of gaps and lower work quality and quantity which is typical for cluster 3. This will be further discussed in Section 5.

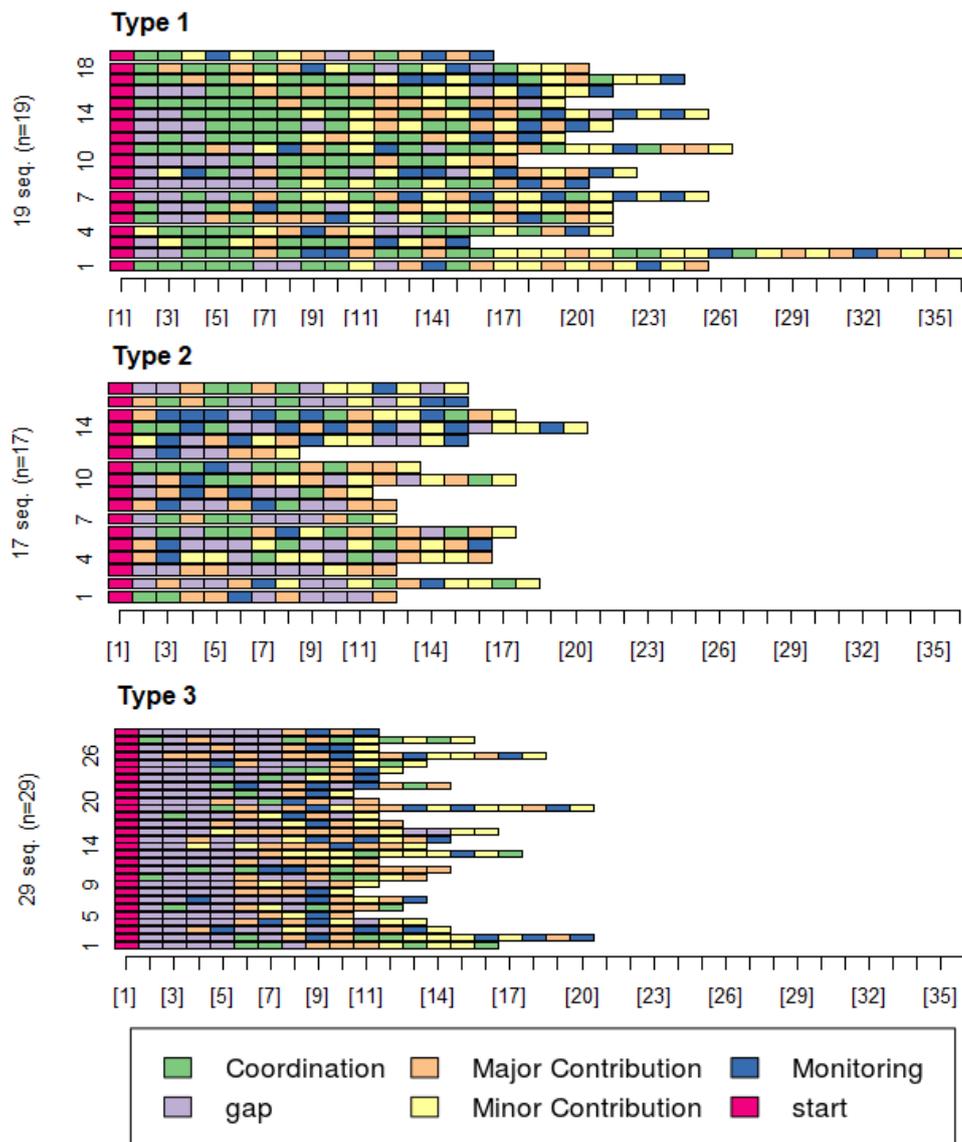


Figure 2. The resulting groups of the PAM clustering.

5. Discussion and Conclusion

The sequence clusters extracted from the dataset show typical distributions of the action types over the whole sequences. The clusters also differ in terms of “productivity” (using the simple word count measure), quality (using concept coverage) and fair work distribution.

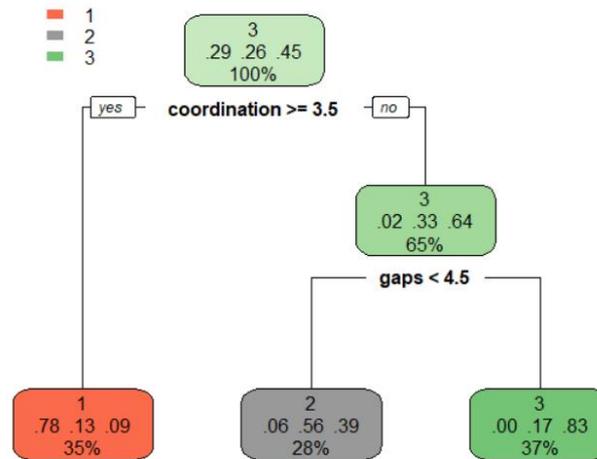


Figure 3. Decision tree describing the 3 clusters.

In the monitoring of groupwork, it is quite intuitive to use observed inactivity as a main indicator for groups being dysfunctional and thus requiring support. Our findings differentiate this assumption: Phases of inactivity (or “gaps”) can be counterbalanced by coordination especially at the beginning of a group task. While the second cluster does not differ much in inactivity from the first one (avg. no. of gaps: 2.7 vs. 3.6), its productivity is almost as low as the one of the third cluster with much higher inactivity (4.8). Indeed, the decision tree analysis indicates that the number of coordination messages differentiates most in terms of the well-functioning of group work. While the productivity of cluster 2 is lower than the productivity of cluster 1, the concept coverage and fair work balance of the clusters are similar. In contrast to that, cluster 3 shows the lowest productivity and concept coverage and the highest work imbalance.

More concretely, it can be said, two days of inactivity in the first half goes along with low productivity in total if no coordination has taken place beforehand. One reason could be that some students are reluctant in making the first contribution. If nobody shows visible activity this creates uncertainty about the motivation and potential for contributing on the part of the groupmates. However, if the groupmates show more “presence” and share their time schedule and structure their group work beforehand, activity gaps are less of an issue and can be overcome. Furthermore, more coordination also results in a fairer work distribution and a higher quality of the produced texts (concept coverage). Again, the difference between cluster 1 and 2 was very low with respect to these measures but very salient between the not well coordinated groups in cluster 3 and the others. On the one hand, these results show that early coordination and presence of activities of group mates is important for distributing the work more equally. On the other hand, a lower work imbalance can also be a result of higher overall engagement of the group members when activities of one member trigger activities of others. The higher concept coverage, and therefore, potentially higher engagement with the course content in better coordinated groups could be a result of the better distribution of work where each member can play a part in specific sub-topics of expertise. Closer examining the relationship between detectable sequence patterns and those collaboration quality indicators is one of the main future research directions.

At present, our findings lead to the following suggestions for establishing and supporting productive group work in large online courses:

- Group assignments should be clearly structured and well defined. Too many degrees of freedom increase the coordination effort for groups which can be a source of problems as our results show. Deadlines have to be adjusted such that students with different time schedules have a chance to bring in their contributions (c.f. Erdmann et al., 2017).
- Scaffolding mechanisms for group work should be offered, for example, guidelines for students highlighting the importance of early coordination. On the technical side, this can be supported by proper communication and scheduling tools.
- The findings presented in this paper can further be used to design intelligent support systems that react on the observation of critical patterns i.e. the absence of coordination in conjunction with

activity gaps to help tutors to turn their attention to those groups. To enable a “real-time” analysis, an automated classification of user activities (see Section 3) is required. A first classification model that predicts the class of activities based on textual properties such as number of written characters, duration of the activity and the tool that was used (forum vs. Etherpad) has been developed which will be used to close the loop from the detection of sequence patterns to adaptive system generated scaffolding and intervention mechanisms in the future.

References

- Abbott, A., Tsay, A. (2000). Sequence analysis and optimal matching methods in sociology review and prospect. *Sociological methods & research* 29, 3-33.
- Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J. (1984). Classification and Regression Trees. Monterey, CA: *Wadsworth & Brooks/Cole Advanced Books & Software*.
- Cornwell, B. (2015). Social Sequence Analysis: Methods and Applications. *Cambridge Univ. Press*, New York.
- Doberstein, D., Hecking, T., Hoppe, H. U. (2017, August). Sequence patterns in small group work within a large online course. In C. Gutwin, S. F. Ochoa, J. Vassileva, & T. Inoue (Eds.), *Collaboration and Technology*, Springer. Paper presented at 23rd International Conference on Collaboration and Technology (CRIWG), Saskatoon, SK, Canada, (pp. 104-117).
- Erdmann, J., Rummelm N., Christmann, N., Elson, M., Hecking, T., Herrmann, T., Hoppe, H.U., Krämer, N., Kyewski, E., Wichmann, A. (2017). Challenges in implementing small group collaboration in large online courses. in *Proc. of International Conference of Computer Supported Collaborative Learning (CSCL)*. Philadelphia, PA, USA, vol. 2, pp. 625–628.
- Ferschke, O., Howley, I., Tomar, G., Yang, D., Rosé, C.P. (2015). Fostering discussion across communication media in massive open online courses. in *Proc. of the 11th International Conference on Computer Supported Collaborative Learning*, Gothenburgh, Sweden, pp. 459-466 (Vol. 2).
- Grünewald, F., Meinel, C., Totschnig, M., Willems, C. (2013). Designing MOOCs for the support of multiple learning styles. in *Proc. of the European Conference on Technology Enhanced Learning*, pp. 371-382.
- Kaufman, L., Rousseeuw, P. (1987). Clustering by means of medoids. in *Statistical Data Analysis Based on the L1-Norm and Related Methods*, Y. Dodge, Ed. North Holland/Elsevier, Amsterdam, pp. pp. 405–416.
- Konert, J., Burlak, D., & Steinmetz, R. (2014). The group formation problem: an algorithmic approach to learning group formation. In *European Conference on Technology Enhanced Learning - EC-TEL 2014* (pp. 221-234). Springer LNCS 8719.
- Kyewski, E., Krämer, N., Christmann, N., Elson, M., Erdmann, J., Hecking, T., Hermann, T., Hoppe, H.U., Rummel, N., Wichmann, A. (2016). Is small group collaboration beneficial in large scale online courses? An investigation of factors influencing satisfaction and performance in GroupMOOCs. in *Proc. of the 12th International Conference of the Learning Sciences (ICLS)*, pp. 918-922.
- Piezou, S.L., Donaldson, R.L. (2005). Online groups and social loafing: Understanding student-group interactions. *Online Journal of Distance Learning Administration*, vol 8, 1.
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, vol. 4 (3), pp. 239-257.
- Roberts, T.L., Lowry, P.B., Sweeney, P.D. (2006). An evaluation of the impact of social presence through group size and the use of collaborative software on group member ‘Voice’ in face-to-face and computer-mediated task groups. *IEEE Transactions on Professional Communication* vol. 49, pp. 28-43.
- Shirvani Boroujeni, M., Dillenbourg, P. (2018). Discovery and temporal analysis of latent study patterns in MOOC interaction sequences. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK 18)*. Sydney, New South Wales, Australia, pp. 206-215.
- Staubitz, T., Pfeiffer, J., Renz, C., Willems, C., Meinel, C. (2015). Collaborative learning in a MOOC environment, in *Proc. of the 8th International Conference of Education, Research and Innovation*, Seville, Spain, pp. 8237-8246.
- Tomar, G.S., Sankaranarayanan, S., Rosé, C.P. (2016). Intelligent Conversational Agents as Facilitators and Coordinators for Group Work in Distributed Learning Environments (MOOCs). *AAAI Spring Symposium at Stanford University*, Palo Alto, CA, USA.
- Weinel, M., Bannert, M., Zumbach, J., Hoppe, N., Malzahn, N. (2011). A closer look on social presence as a causing factor in computer-mediated collaboration. *Computers in Human Behavior*, vol. 27, 513-521.
- Wichmann, A., Hecking, T., Elson, M., Christmann, N., Herrmann, T., Hoppe, H.U. (2016). Group formation for small-group learning: Are heterogeneous groups more productive?. in *Proc. of the International Symposium on Open Collaboration (OpenSym)*.