Resource Access Patterns in Exam Preparation Activities

Sabrina Ziebarth, Irene-Angelica Chounta, and H. Ulrich Hoppe

University of Duisburg-Essen, Germany
{ziebarth, chounta, hoppe}@collide.info,
WWW home page: http://www.collide.info/

Abstract. In this paper we investigate patterns of resource usage for exam preparation based on a resource intensive blended learning course. To this end, we analyzed a blended learning, online course facilitated by Moodle. During the course, the students had to work individually and in teams. Furthermore, they were given access to a broad spectrum of learning resources such as videos, slides, wiki articles and quizzes. The logfiles obtained from Moodle were further processed and analyzed. Our analysis approach is based on association rule as well as sequential pattern mining. The results indicate that students' activity with respect to resource usage follows common patterns during exam preparation either on the individual or the group level. These patterns also relate to the performance of students and to reflect their prior collaborative experience.

Keywords: Small Private Online Courses, Moodle, Exam Preparation, Sequential Pattern Mining, Association Rule Mining

1 Introduction

There are numerous examples of enhancing traditional learning scenarios with online and interactive elements, usually subsumed under the notion of "blended learning". Armando Fox [5] argues for complementing traditional university courses with concepts and materials from Massive Open Online Courses (MOOCs) to "increase instructor leverage, student throughput, student mastery, and student engagement". He calls this concept "Small Private Online Course" (SPOCs). To investigate the effectiveness of complementing traditional university courses with SPOCs, in 2013 we redesigned an existing Master level lecture on user/task modeling and models of interactive and collaborative learning environments [13]. The main idea was to shift the focus of the classroom experience from teacher centered lecturing to student centered interactions. To achieve that, the lecture was redefined in the sense of providing orientation knowledge ("advance organizer") and core definitions with a limited number of characteristic examples. The presence based exercise was "virtualized", i.e. replaced by a combination of individual and group assignments with reporting back into the learning management system (a Moodle 2 plattform). Additionally, the plenary face-to-face

setting of the lecture was used as a forum to take up and discuss results from the virtual exercise activities. An important assignment was the collaborative creation of wiki articles. Each group focused on another topic, thus the combination of all wiki articles also provided the course with a collaboratively created script of the lecture for exam preparation. Based on the experiences of the 2013 course, in 2014 we repeated the course with some changes, e.g. regarding group formation and size, and added exercises involving virtual labs for supporting the better understanding of complex, theoretical concepts.

This kind of blended learning course results in a broad spectrum of learning resources (lecture slides and videos, papers, wiki articles and quizzes). First results of analyzing the Moodle log files of the 2013 course show that during lecture time the resource usage depends highly on the exercises, while there are different patterns of resource usage in the exam preparation phase [7]. Thus, to analyze how the different types of resources are used for learning we focus on the exam preparation phase. Furthermore, the students of the 2013 course stated that they regarded both teacher as well as student created material as beneficial for learning (with a slight preference for teacher generated content) [13]. Therefore, we want to analyze to which extent this is reflected in their resource access patterns during exam preparation. While this paper only focuses on the log analysis of the Moodle data, we plan to survey additional information in the next instance of the course to allow for triangulation of the results. Our research questions are:

- Are there typical patterns in resource usage for exam preparation?
- Do members of groups portray similar behavior with respect to resource usage during the exam preparation phase and is this reflected on the overall results?

2 Related Work

The analysis of learning activities aims to identify the meaningful episodes that take place when a group of learners interact with the instructor, their peers or the available resources within a learning context. The use of computers allows the detailed recording of user activity in log files and provides access to vast resources for further analysis [3]. The analysis of learning activities usually focuses on user activity and user interaction [12]. However, we can gain valuable information not only from the way users interact with each other but also from the way they interact with the learning material. To that end, Learning Analytics can be applied in order to investigate the relations between learners and learning resources. The use of learning resources in MOOCs and online communities has been widely studied in order to define common learning styles [2] and contributor's profiles for user-created learning material [6]. There is also research demonstrating dependencies between the LMS usage in blended learning courses and the achieved learning results, e.g. [9]. Data-mining algorithms are widely applied in order to discover patterns on resource usage and consequently for further assessment of good practices [11]. Resource usage has been proposed

as a method for clustering learning objects in a way that reflects their semantic similarity [10].

3 Methodology

As basis for this analysis the Moodle log files of the course, that was conducted in summer term 2014, have been used. 33 students participated in the course, but two deleted their accounts and thus were not considered. The students formed eight groups of four students on average and these groups were maintained over the complete lecture time. Twenty-six students took their exam. The exams were conducted orally over a time span of one week, starting 18 days after the last lecture. In this analysis we focus on these 18 days as well as the exam week. Of course, only the behavior on the Moodle platform can be observed. Students could easily download lecture slides and scientific papers, thus they would not need to go back to the platform. However, it was much harder to download the lecture videos, since they were provided in a player that did not offer downloads. Furthermore, Moodle does not offer a download function for wiki articles and quizzes could only be attempted in the Moodle platform.

3.1 Association Rule Analysis

This section focuses on the research question regarding patterns in resource usage for exam preparation. We consider resources used together in one Moodle session to reflect the student's learning goal of that session (and thus a part of his/her learning strategy). Such a goal could be to recapitulate a certain subtopic using different types of relevant resources, to recapitulate a broader topic using only specific types of resources or to use quizzes or slides with exam questions for self-testing. Based on this assumption, resources that often occur together in one session can be considered to be associated in the students learning strategies.

To detect associations between resources based on their co-occurrence in sessions, the Apriori algorithm [1] was applied. Table 1 shows examples for association rules found in the Moodle sessions of the exam preparation phase. The rules with the highest support are between resources of the same media type and of the same topic. All of the 68 rules with a minimal support of 0.05 and a minimal confidence of 0.5 consist of resources that belong to the same topic. Most of them only refer to slides, but there are also four rules on wiki articles, two rules on videos and two referring to one slide set and one of the papers highlighted for exam preparation.

The Apriori algorithm can also be used to create (undirected) associations between resources that are weighted by the support value ("co-occurrences"). The results of analyzing the 185 co-occurrences with a minimal support of 3 % of learning objects support the observation that students mainly use resources of the same media type and topic in one session. But while there are some interconnections between different types of teacher generated content (slides, papers, videos), the student generated wiki articles are only linked to other wiki

articles. This indicates that while wiki articles are used for exam preparation, they take another role than the teacher generated material. While there are typically few links between resources on different topics, the accessed lecture slides are more often linked with slides of other topics. This could e.g. be an indicator for downloading slides to have a local copy, getting an overview or searching for some specific information. Reducing the minimal support to 2 % shows a triad consisting of the three quizzes. This indicates that the quizzes are used together, but separately from the other resources. Since they cover the three main topics of the lecture, there is no connection by topic. A typical goal for using the three question might be self-testing, which fits with the observation that quizzes are only used in the second part of the exam preparation phase.

Table 1. Examples for Association Rules

Rule	support	confidence
Slides: Cognitive Task Models Q&A-2 → Slides: Cognitive (Task) Models Q&A	0.085	0.848
Paper: Understanding Interactivity → Slides: Interactivity, Direct Manipulation and User Centered System Design	0.057	0.613
Wiki article: GOMS → Wiki article: Model Human Processor	0.051	0.739

3.2 Sequential Pattern Mining Analysis

Learners with common interests are expected to interact with thematically-related learning material [8]. We aim to explore the sequential patterns of resource usage among the members of student groups. The research hypothesis is that groups who collaborated successfully during a course, will also portray similar behaviour during the exam preparations with respect to resource usage. Moreover, we explore the relation of resource usage and the learning outcome as assessed by the teacher of the course. For the purpose of our study, we portrayed the activity of groups as sets of sequences. Each sequence maps the activity of individual users with respect to resource usage. A sequence is composed of itemsets that represent the resources accessed within one user session. According to the hypothesis, successful collaborative practice should lead to similar patterns of resource usage. Therefore, the activity of each student is compared to the activity of his team mates. To discover potential frequent sequential patterns in user activity, we used the PrefixSpan algorithm [4] and looked for patterns that occur for all the students of one group (100% pattern matching). We argue that frequent patterns of resource usage characterize groups that share common ground and have adopted common working styles. The PrefixSpan algorithm discovered frequent patterns for all eight groups. This means that all members per group accessed the same resources in the same sequence. The number of sequential patterns ranges from 2 to 45 and their length ranges from 1 to 3 itemsets (2 to 13 objects). Group 1 scored the lowest average group grade and the

frequent sequential patterns consisted only of video resources. The same is true for groups 3 and 4; group members accessed in common only wiki articles. For Group 6, that scored the higher group grade, the frequent patterns of resource usage consisted of 6 objects from four different categories: two scientific papers, one lecture slides, one video and 2 wiki articles. The group grade appears to correlate with the number of categories of resources. Groups that use resources from more categories also score higher grades, thus students who spread over various resources achieve better results in the exam. However, the overall number of accessed resources is similar for all groups and there is no indication that groups with higher grades access more resources in general. To further study whether user behavior in exam preparation is affected by group dynamics, we divided groups in two categories: heterogeneous and homogeneous. We define homogeneity (hf) as the deviation of the individual highest grade within the group to the group's average grade. Heterogeneous groups (Group 3 with hf = 0.87 and Group 4 with hf = 2.68) portrayed the minimum number of frequent patterns on a 100% level (i.e. for all the users). For these two groups, there were indeed two students (one per group) that had limited activity in comparison to their team members. Thus, it was difficult to discover sequential patterns that would appear in all users activity. Additionally, homogeneous groups have more sequential patterns that consist of more resources from various categories. This pinpoints that students of the same level access similar resources in the same way. Furthermore, students of different levels use different kinds of resources. This consequently can serve as an indication that particular group members do not share a common ground with the rest of the group. However, in the current study we could not establish statistical significance or relevance due to the small size of the dataset. Therefore, further exploration is necessary in order to establish general conclusions.

4 Conclusions

In this paper we study the activity of students during the exam preparation phase of a resource-intensive, blended-learning course. The course was supported by an LMS platform (Moodle) and a wide variety of learning resources was provided to and partially created by students. We assume that the resource usage will reflect the practice of students, thus providing valuable information on how students prepare for the exams. To test this assumption we are applying questionnaires regarding the usage of sources and exam preparation strategies in the current summer term and plan to triangulate the results with the log analysis. We hope to use the detected practices to predict success of and recommend resource usage strategies for students in future. The analysis of student activity with respect to resource usage was carried out on both the individual and the group level. Our main objective was to explore whether there are typical patterns in resource usage during the exam preparation phase and if the members of the same group portray similar behaviour regarding resource usage. To that end we used association rule analysis and sequential pattern mining. The results indicate there are

common patterns of resource usage on both the individual and the group level that relate to the performance of students, to the type of resources and to their prior collaborative experience. However, it is necessary to conduct further, large scale studies in order to provide well-established conclusions. Additionally these studies should not only address log analysis but also surveys of the students (e.g. regarding learning attitudes and strategies, motivation, group interaction, etc.) to complement the results and to allow for triangulation.

References

1. Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., Verkamo, A.I.: Knowledge Discovery and Data Mining, chap. 12, pp. 307–328. AAI Press / The MIT Press (1996)
2. Breslow, L., Pritchard, D.E., DeBoer, J., Stump, G.S., Ho, A.D., Seaton, D.: Studying learning in the worldwide classroom: Research into edxs first mooc. Research & Practice in Assessment 8, 13–25 (2013)
3. Dillenbourg, P., Baker, M.J., Blaye, A., O'Malley, C., et al.: The evolution of research on collaborative learning. Learning in Humans and Machine: Towards an interdisciplinary learning science. pp. 189–211 (1995)
4. Fournier-Viger, P., Gomariz, A., Soltani, A., Lam, H., Gueniche, T.: Spmf: Open-source data mining platform. `http://www.philippe-fournier-viger.com/spmf/` (2014)
5. Fox, A.: From moocs to spocs. Communications of the ACM 56(12), 38–40 (2013)
6. Halatchliyski, I., Hecking, T., Goehnert, T., Hoppe, H.U.: Analyzing the path of ideas and activity of contributors in an open learning community. Journal of Learning Analytics 1(2), 72–93 (2014)
7. Hecking, T., Ziebarth, S., Hoppe, H.U.: Analysis of dynamic resource access patterns in a blended learning course. In: The 4th International Conference on Learning Analytics and Knowledge. pp. 173–182. ACM New York (2014)
8. Hoppe, U., Pinkwart, N., Oelinger, M., Zeini, S., Verdejo, F., Barros, B., Mayorga, J.I.: Building bridges within learning communities through ontologies and thematic objects. In: Proceedings of the International Conference on Computer Supported Collaborative Learning. pp. 211–220 (2005)
9. Mödritscher, F., Andergassen, M., Neumann, G.: Dependencies between e-learning usage patterns and learning results. In: Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies. ACM (2013)
10. Niemann, K., Schmitz, H.C., Kirschenmann, U., Wolpers, M., Schmidt, A., Krones, T.: Clustering by usage: higher order co-occurrences of learning objects. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. pp. 238–247. ACM (2012)
11. Perera, D., Kay, J., Koprinska, I., Yacef, K., Zaïane, O.R.: Clustering and sequential pattern mining of online collaborative learning data. Knowledge and Data Engineering, IEEE Transactions on 21(6), 759–772 (2009)
12. Suthers, D.D., Dwyer, N., Medina, R., Vatrapu, R.: A framework for conceptualizing, representing, and analyzing distributed interaction. International Journal of Computer-Supported Collaborative Learning 5(1), 5–42 (2010)
13. Ziebarth, S., Hoppe, H.U.: Moodle4spoc: A resource-intensive blended learning course. In: 9th European Conference on Technology Enhanced Learning (EC-TEL 2014). pp. 359–372. Springer International Publishing, Graz, Austria (2014)