5-24-2013

Traces of Students' Performances in Online Activity Logs

Elahe Javadi
Illinois State University, ejavadi@ilstu.edu

Thanesh Rajandran
Illinois State University, tprajan@ilstu.edu

Traces of Students' Performances in Online Activity Logs

Elahe Javadi
Illinois State University
ejavadi@ilstu.edu

Thanesh Rajandran
Illinois State University
tprajan@ilstu.edu

ABSTRACT

In this study we examine the question of whether or not students' activities on an online learning content management systems (LCMS) can be an indicator for students' performances in the course. We measure performance by grades in three different exams and we use variety of measures for quantifying the online activities. The data is collected from two sections of a two hundred level information systems course in a Midwestern university. Our data analysis results are partially consistent with findings of prior literature in this area that suggests importance of students' consistent access to the course material throughout the semester. Our mixed results, however, calls for a closer examination of the measures that have been developed and used in the literature and their effectiveness in inferring students' performance based on students' online activities. The analyses presented here and similar models can help instructors plan course-related actions and implement interventions after close examinations of the data that is freely available to them through online learning content management systems.

KEYWORDS
Student performance, online activities, access, consistency

INTRODUCTION

Each semester, learning content management systems collect vast amount of data about students' online activities. The data is available for instructors to view and analyze, typically in pre-defined and relatively limited formats and structures (e.g., bar charts, time series data, and summary statistics). The instructors are not generally equipped with tools to perform deep data explorations to extract instructional knowledge from the raw data that is available to them.

This study's premise is that online activity logs can be transformed to actionable knowledge for instructors and students. We believe that LCMS data can be mined for discovering useful models and patterns that can help instructors and students better assess their respective roles and performances in achieving their educational goals. The mined models can provide easy-to-understand and easy-to-use educational guidelines for students and instructors. Educational guidelines will provide students with new ways to identify their strength and also areas for improvement. It can also shed light on otherwise neglected performance glitches and can therefore encourage outside-classroom effort. For instance, analysis could inform students about their performance compared to their classmates and map this comparison on the extent of their online activities. Instructors can use the data-driven evidence to change or fine-tune their course design and delivery methods. For instance, instructors can devise and implement in-semester interventions. Interventions can be in variety of forms such as simple reminders or data-driven review meetings and one-on-one follow-ups with students. In other cases, LCMS data may suggest that consistent use of certain online resources (Biktimirov & Klassen 2008), such as practice quizzes, enhanced students' performance in their exam, but the night-before-the-exam rush to the same resource did not have any effect on students' performance in that exam; here, the instructor can use in-classroom techniques to reinforce the importance of over-semester use of practice quizzes. Data-driven evidence can answer a variety of other questions such as:

1. Does resource-use lead to better performance in the course?
2. Which is more important, frequency or consistency of resource-use (Baugher *et al.* 2003)?
3. Which resources are better predictors for performance in class?
4. Has using a resource triggered accessing other resources?
5. If resource-use is essential for success, can we identify in-danger segments of student?
6. Are some resource segments used more frequently and more consistently than others? What are their common characteristics?
7. Are there any distinctive patterns based on the course level? (e.g., resource use in 100-level courses is not strongly linked to performance)
8. Are there any trends when majors, schools/departments, or colleges are compared?

Finding answers to the above questions can be prescriptive in the selection and use of LCMS features for particular types of classes or major; students can also identify which LCMS features work best for their learning style, and higher education institutions can better assess how the LCMS can be tailored to their programs. Prior research studies have examined LCMS from different perspectives. An overview of prior research is provided in the following section.

BUSINESS INTELLIGENCE IN EDUCATION- BACKGROUND

STUDENT INVOLVEMENT & STUDENT PERFORMANCE

Online activities can be analyzed for the quantity, quality, patterns and consistency of those patterns during the semester. To use students' online activities as a basis for inferring students' performance, researchers have studied a variety of measures. The first measure is the volume or number of resource hits by students (Asarta and Schmidt 2013; Baugher *et al.* 2003; Biktimirov & Klassen 2008). Researchers have also looked at what they name, intensity, which applies the concept of diminishing returns to the sheer number of access (Asarta and Schmidt 2013). The premise of intensity measure is that the earlier instances of access to the course material are more effective that the latter ones.

Because the total hit has not been found to be a powerful predictor of performance, the measure has been refined to represent number of unique files that the students accessed (Biktimirov & Klassen 2008) or number of the unique files that were accessed by students over the total number of files (completeness) (Asarta and Schmidt 2013).

Alternatively, researchers have looked at the formative component of volume which includes access to different resource groups such as homework, homework solutions, PowerPoint slides, or exam solutions (Biktimirov & Klassen 2008). It should be noted that based on idiosyncratic features of a course design and final grade computation in the course, access to specific group of resources may be found to be correlated with performance. For instance, Biktimirov and Klassen (2008) found that access to homework solutions was highly correlated with performance.

Timing and consistency of access to the course material are two other measures that researchers have examined. Consistency is measured in different ways. Researchers have measured consistency by whether or not a student accessed (any) course material during a week (Biktimirov & Klassen 2008) or between class meetings (Baugher *et al.* 2003). In their research, Baugher *et al.* (2003) found that consistency was significantly correlated with the performance. Timing of access is an innovative measure that extracts details of students approach to use of online course material. Asarta and Schmidt (2013) created unique measures of timing of access to represent pace of students' access, students' tendency to avoid cramming and students' reviewing style; the researchers then found timing to be a significant predictor of student performance.

LEARNING CONTENT MANAGEMENT SYSTEMS& ANALYTICS

Many academic institutions are beginning to see the benefit of leveraging their data to analyze a wide range of factors that impact how they can make better decisions (Information Management 2013). Learning Content Management Systems provide one source of data that academic institutions can leverage. The LCMS technology is an important piece of students' success in an online or hybrid environment (Volery 2000). Large amounts of data are available from the LCMS that can help researchers map student activities in the LCMS to the learning process (Martín-Blas and Serrano-Fernádez 2008). This study builds on the work of Romero *et al.* (2008) to extend and apply to the Sakai LCMS environment.

Based on the involvement and performance studies identified above, there is the opportunity to integrate many components of Business Intelligence and Analytics to examine the impact of the LCMS on student performance. The study proposed here uses the visualization approach identified by Mazza and Dimitrova (2004) to an open source installation at a large academic institution. The data provided in the Sakai will allow us to study in depth the impact on the types of assignments or features used (i.e., discussion thread, web based assignment, blog, etc.) within different learning environments (ie, face-to-face, hybrid or online) as seen in previous work (Bonham *et al.* 2003, Larson and Sung 2009).

Next section summarizes findings of a project that has been completed on two sections of a two-hundred level information systems course in a Midwestern university.

DATA COLLECTION

The current study is performed on the data collected from two sections of a 200-level course in the area of Information Technology. In this section data is examined from different perspectives. In the analysis, we focused on frequency and consistency of access to a specific group of resources called *Reading Assignments*, and we examined the correlation among the two measures and students' performance in class (final grade) and three separate examinations. Reading assignments in

this course were a group of resources that helped students with their readings and preparations before the class. Each reading assignment included up to 7 questions and up to 7 terms. To answer the questions on the reading assignments, students could either read the suggested book chapters from the text book, or use other books or online resources (also suggested in the reading assignments). Students were not asked to return the completed assignments but working on the reading assignments would help students better understand the upcoming lectures and would enable students to actively participation in the class discussions. There were 15 reading assignments during the semester.

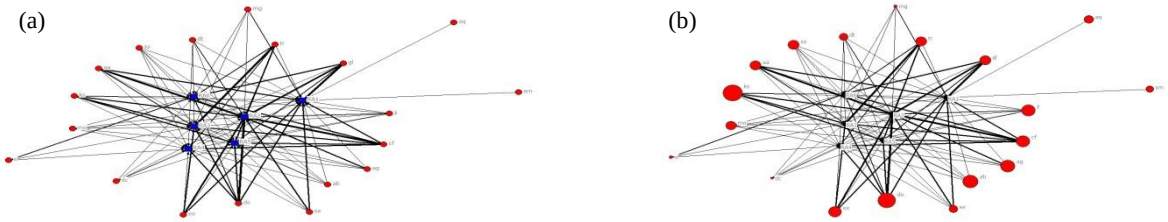

Figure 1: Network of Students & Resources

Figure 1.a shows resource use for students in the pilot study's class section 1. Nodes in the center depict resources and the red nodes represent students. Link thickness represents frequency of resource use by each student. Figure 1.b shows resource use and its relation to performance in course's first examination. Node sizes are proportionate to students' grades on the exam. A close examination of this network diagram shows that some resources have been better predictors of the performance than others.

DATA PROFILE

For further analysis, access to reading assignments 1-15 was combined to create measures of hit frequency and consistency. Hit frequency was the number of access instances to all reading assignments over 17 weeks (16-week semester plus the finals week). Hit consistency was measured by consistency over 17 weeks. If RAs were accessed during a certain week, we flagged it. The number of flagged weeks is consistency. Figure 2, shows frequency of visits by all students over the semester.

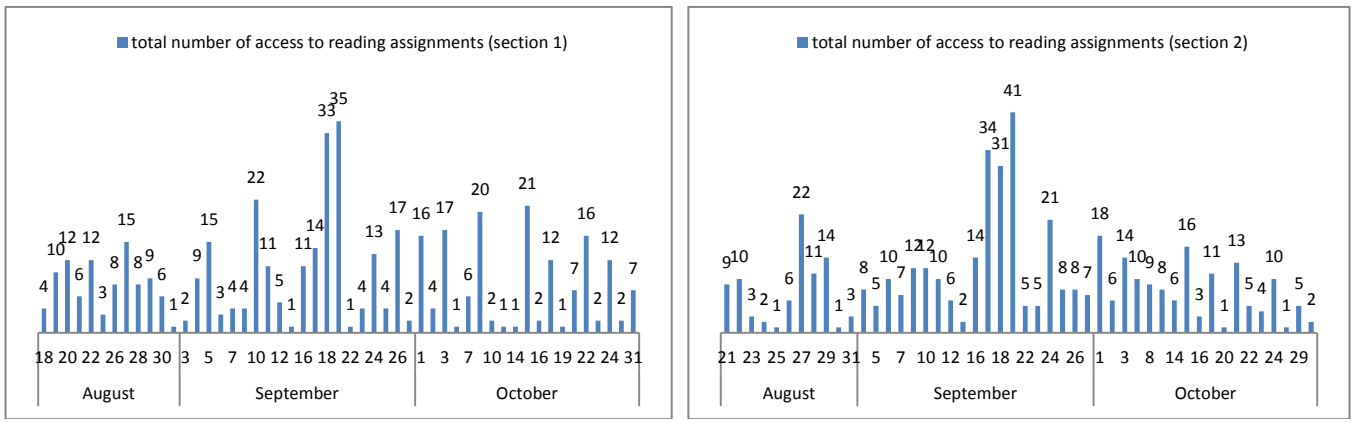

Figure 2: Total Hits by All Students during the Semester

The data above was then decomposed to show each student's hit frequency and consistency over the 17 weeks. Table 1 depicts the data table for section 1 in the pilot study.

Table 1: Access Numbers by Student by Week

SID	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11	w12	w13	w14	w15	w16	w17
S1	2	2	0	0	3	2	2	1	2	1	0	0	0	0	0	0	2
S2	4	4	0	4	3	3	2	1	2	1	0	0	0	0	0	0	4
S3	0	3	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0
S4	0	4	0	0	1	0	2	1	1	1	0	0	0	0	0	0	0
S5	4	6	1	2	2	4	3	3	4	3	0	0	0	0	0	0	4

S6	1	2	3	2	13	2	2	1	2	6	0	0	0	0	0	0	1
S7	0	3	1	1	1	2	6	1	1	1	0	0	0	0	0	0	0
S8	3	0	4	0	0	0	0	0	0	9	0	0	0	0	0	0	3
S9	2	2	5	0	15	1	0	0	2	7	0	0	0	0	0	0	2
S10	0	2	0	0	6	1	1	1	2	0	5	0	0	0	0	0	0
S11	0	0	0	5	6	0	0	1	2	1	0	0	0	0	0	0	0
S12	5	0	2	1	1	2	2	1	2	1	0	0	0	0	0	0	5
S13	3	2	2	4	2	2	3	1	2	1	0	0	0	0	0	0	3
S14	4	2	1	1	2	1	2	1	2	1	0	0	0	0	0	0	4
S15	0	0	0	1	7	1	1	1	1	1	0	0	0	0	0	0	0
S16	0	0	0	1	7	1	1	1	1	1	0	0	0	0	0	0	0
S17	0	6	1	0	7	1	1	2	1	0	1	0	0	0	0	0	0
S18	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1
S19	0	1	0	0	0	1	2	1	1	1	1	0	0	0	0	0	0
S20	0	0	2	0	1	2	2	2	2	0	0	0	0	0	0	0	0
S21	2	0	5	3	2	2	2	1	2	0	0	0	0	0	0	0	2

Table 2: Frequency and Consistency Measures for Students

Table 2 includes some descriptive information about access frequency and consistency for students over the 17 weeks period.

	Frequency(over 17 weeks)	**Consistency (over 17 weeks)**
Section 1 (N=21**)**	Min=4, Avg.=19, Max=36	Min=2, Avg.=7.7, Max=11
Section 2 (N=25)	Min=1, Avg.=16.3, Max=35	Min=1, Avg.=6.8, Max=12

Figure 3.a and 3.b are created based upon Section 1 data and depict students' access to a specific group of course material called Reading Assignments. The nodes represent students. In Figure 3.a and Figure 3.b, link thickness represents frequency of resource use by each student based on frequency of resource use over a 17-week period and consistency of resource use over the same period. Figure 3.c and 3.d illustrate the same information for Section 2 of the course. A close examination of the two network diagrams shows that consistency of resource use is slightly better predictor of performance in class than frequency of resource use, Pearson correlation 0.41(hit consistency & course performance) vs. 0.35 (hit frequency & course performance) in Section 1 and 0.2 (hit consistency & course performance) vs. 0.1(hit frequency & course performance) in Section 2.

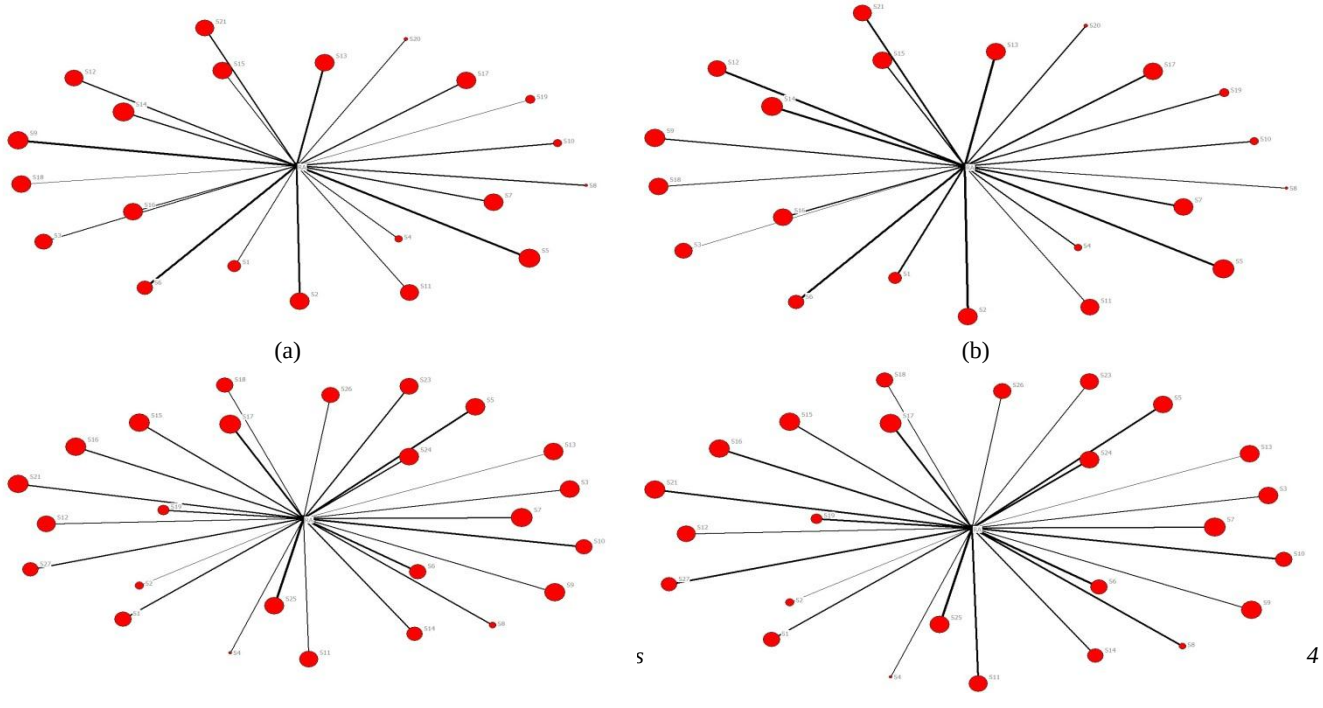

Figure 3: Access Frequency and Consistency & Course Performance

Because there were three examinations in the class, we replicated the analysis for the 3 examinations (not the overall performance in the course). The results are shown in Table 3.

Table 3: Pearson Correlations for Different Time Periods

Pearson Correlations	**Frequency w1-w5**	**Consistency w1-w5**	**Frequency w6-w9**	**Consistency w6-w9**	**Frequency w1-w17**	**Consistency w1-w17**
Section 1	**Exam 1:** 0.62	0.67	**Exam2:** 0.33	0.24	**FinalExam:** 0.54	0.58
Section 2	-0.04	-0.04	0.1	0.8	0.34	0.21

The above numbers call for refinement on the consistency measure or a measure that would take into account intensity and consistency.

CONCLUSION

The plan for this research study is to collect data from different courses across campus in order to discover trends and patterns that hold true in more than one class or for more than one instructor. Patterns and trends discovered in data can answer the questions listed in the introduction. It is anticipated that this study's analyses will find patterns and trends that are currently hidden and thus cannot be used or acted upon by instructors or students. Finding descriptive and predictive models will provide new methods for achieving educational excellence in classes.

REFERENCES

Asarta, C. J. & Schmidt J. R. (2013). Access Patterns of Online Materials in a Blended Course. *Decision Sciences Journal of Innovative Education*. (11:1).

Baugher, D., Varanelli, A., & Weisbord, E. (2003). Student hits in an internet-supported course: How can instructors use them and what do they mean? *Decision Sciences Journal of Innovative Education.* (1), 159–179.

Biktimirov, E. N. and Klassen K. J. (2008). Relationship between use of online support materials and student performance in an introductory finance course. *Journal of Education for Business*. (83), 153–158

Bonham, S. W., Deardorff, D. L., & Beicher, R. J. (2003). Comparison of Student Performance Using Web and Paper-Based Homework in College-Level Physics. *Journal of Research in Science Teaching.* (40:10), 1050–1071.

Borgatti, S.P., Everett, M.G. and Freeman, L.C. (2002). Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

Infromation Management (2013). 8 Institutions Bringing Value to BI Education. *Information and Management*. http://www.information-management.com/gallery/BI-information-management-jobs-college-university-10021400-1.html (accessed March 1, 2013).

Larson, D. K. and Sung, C-H (2009). Comparing Student Performance: Online Versus Blended Versus Face-to-Face. *Journal of Asynchronous Learning Networks.* (13:1), 31-42.

Mazza, R. and Dimitrova, V. (2004). Visualising Student Tracking Data to Support Instructors in Web-Based Distance Education. *13th International World Wide Web Conference (WWW 2004) - Alternate Educational Track*. May 17-22. 154-161.

Martín-Blas, T. and Serrano-Fernádez, A. (2008). The role of new technologies in the learning process: Moodle as a teaching tool in Physics. *Computers & Education*. (52), 35-44.

Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*. (51:1), 368-384.

Volery, T. (2000). Critical Success Factors in Online Education. *The International Journal of Educational Management.* (14:5), 216-223.