Internet and Higher Education 11 (2008) 65–70

Contents lists available at ScienceDirect

Internet and Higher Education

Data for free: Using LMS activity logs to measure community in online courses

Erik W. Black *, Kara Dawson, Jason Priem

University of Florida, College of Education, School of Teaching and Learning, United States

ARTICLE INFO

Article history:
Accepted 17 March 2008

Keywords:
Community
Learning management systems
Data logs
Prediction

ABSTRACT

In the study of online learning community, many investigators have turned attention to automatically logged web data. This study aims to further this work by seeking to determine whether logs of student activity within online graduate level courses related to student perceptions of course community. Researchers utilized the data logging features of the Moodle learning management system and the Classroom Sense of Community Index. Results reveal that cumulative course data logs are predictive of both a student's sense of connectedness and student community. This study adds to a foundation for a non-invasive assessment of affective variables in online learning environments, and suggests a simple method for providing e-learning instructors with real-time feedback for fostering online community.

1. Introduction

The United States has experienced unprecedented growth in both the availability of and participation in online education programs. At present, there are over 3.2 million online students at the college and university levels (Allen & Seaman, 2007) and over 96% of the very largest higher education institutions have online course offerings (Allen & Seaman, 2006). Over 700,000 students also participate in K-12 online education (Smith, Clark, & Blomeyer, 2005). Clearly, online learning is experiencing phenomenal growth. However, methods and tools for researching experiences within online communities have not kept pace; education lags behind industry and government in the use of comprehensively-gathered and carefully-analyzed data to support decision making (Black, Ferdig, & DiPietro, 2008). The growing use of learning management systems (LMS), many of which automatically keep logs of student activity, presents an exciting means of narrowing this gap. Lately, many researchers have worked to exploit this potential, both in academic research and the design of practical online learning applications. The present study continues this work, seeking to explore whether students' perceptions of community can be measured via logs of student activity within graduate level online courses. Since feelings of community are known to significantly affect online learning performance, such a simple and immediately accessible measurement of this affective variable would be useful for online learning instructors and researchers alike. Further, the measurement would provide a non-invasive alternative to currently-employed survey methodologies. This is a growing need as students increasingly develop "survey fatigue," apathy toward completing surveys. This paper will begin with a discussion of the importance of community in online learning.

2. Literature review

2.1. Community in online learning

Throughout the last 10–15 years, online learning researchers and instructional professionals have promoted the significance of community in online learning environments (Wallace, 2003).This importance is likely only to grow as online students increasingly come to see community as a fundamental part of online life (Weller, 2007). Collaboration between both students and online teachers is necessary to effectively cultivate a thriving online community (Berge & Collins, 1995; Palloff & Pratt, 1999). According to Wallace (2003), community in online environments arises at the intersection of three contemporary components in educational research: social learning theories, the affordances of computers as communication devices and increased utilization of theory in online course development.

Rovai (2002c) defines community in online learning environments as:

> …consisting of two components: feelings of connectedness among community members and commonality of learning expectations and goals….Classroom community is strong when learners (a) feel connected to each other and to the instructor, (b) manifest the immediate communication behaviors that reduce social and psychological distance between people, (c) share common interests and values, (d) trust and help each other, (e) actively engage in two-way communications, and (f) pursue common learning objectives. (p. 322)

* Corresponding author.
E-mail address: erikwblack@gmail.com (E.W. Black).

1096-7516/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.iheduc.2008.03.002

Hung and Chen's (2001, p.10) dimensions of principles of learning support Rovai's definition of online community:

1. Situatedness: fostered by contextualized activities, e.g. tasks and projects based on demand and needs.
2. Commonality: fostered by shared interests, e.g. in books; and shared problems.
3. Interdependency: fostered by varying expertise levels; varying perspectives or opinions; varying needs, mutual benefits; and complementary motives.

This further grounds the concept of online community within the work of Vygotsky and Spiro.

It is clear that community is an essential part of successful online education. Limited face-to-face communication can lead to feelings of isolation which, in turn, can lead to dissatisfaction, poor performance and course non-completion (Cereijo, Young, & Wilhelm, 2001; Curry, 2000; Rovai & Wighting, 2005). Research by Haythornthwaite, Kazmer, Robins, and Shoemaker (2000) relates feelings of isolation to a low sense of community. Findings by Eastmond (1995) indicate that isolation can be alleviated when learners support one another. Additionally, Rovai (2002b) has demonstrated that encouraging a sense of community will effect student satisfaction, learning and retention.

Given this well-established importance of community in online learning, instructors and administrators are typically keen to foster a sense of community in online learning students (Mazzolini & Maddison, 2007). However, the nature of online learning often makes this troublesome. Specifically, the lack of face-to-face interactions in the online environment makes it very difficult to appraise online classroom community (Vrasidas, 2004; Mazza & Milani, 2005; Mazzolini & Maddison, 2007). Lacking access to the same breadth of social indicators as their classroom counterparts, online learning instructors must assess community through a diminished interaction "bandwidth" (Van Lehn, 1988). It is little surprise, then, that instructors are often mistaken in their assessments of online social situations such as class discussions (Mazzolini & Maddison, 2007).

Researchers, administrators, and instructors have turned to survey data to answer questions relating to classroom community (Rovai & Wighting, 2005). However, there are significant limitations to this approach. First, today's online students are over-surveyed (Dillman, 2002), subjected to increasing numbers of surveys and assessments seeking to understand their motivations, concerns and mind-set. Students see little relevance in many of these surveys, increasing student apathy and non-response (Kalton, 2000; LaBruna & Rathod, 2005). Some universities, recognizing that "…student cooperation with surveys [is] a scarce and valuable resource that should be used wisely," have begun to institute policies guiding and limiting survey access to students (Porter, 2005). This one–two punch of decreasing reliability and availability of survey data will no doubt impact the usefulness of this methodology. Second, assessment tools, such as surveys, necessary for the measurement and evaluation of key factors that equate to online learning success have not kept pace with online education's explosive growth. A limited range of assessments are available for use within online education programs and few of these have proven valid and reliable (Black et al., 2008).

2.2. Non-invasive measures in online environments

In order to satisfy the need for valid and reliable assessment tools in today's environment of survey-saturated students, many have advocated adopting new approaches to data-gathering (Sinickas, 2007; Gofton, 1999). Until recently, educators seemed reticent to embrace data mining and statistical techniques to analyze data recorded by computing media themselves (Lopes & David, 2006; Lowes, Lin, & Wang, 2007; Klassen & Smith, 2004); however, such methods are now rapidly gaining popularity (Romero & Ventura, 2007). A common theme to these approaches is that they are less intrusive and subjective, though typically requiring more processing than survey methods (Pahl, 2004). Within this general paradigm of non-invasive assessment several different approaches have emerged, each with its own advantages and weaknesses. Researchers have made use of data from three main sources: (1) recorded text, (2) web server log files, and (3) learning software log files. Several such studies are listed in Table 1.

2.3. Recorded text

Several authors (Dringus & Ellis, 2005; Lowes et al., 2007; Mazzolini & Maddison, 2007) have employed data mining of text communications in learning management systems (LMS) and computer supported collaborative learning (CSCL). This is a particularly rich source of data which has yielded significant findings. Unfortunately, while automated text mining using artificial intelligence algorithms has shown considerable promise in educational applications (Mochizuki et al., 2005; Tane, Schmitz, & Stumme, 2004), mining for relatively subtle social indicators remains impractical (Dringus & Ellis, 2005). Consequently, this methodology is limited by the need to perform relatively labor-intensive hand-coding.

2.4. Web server log files

Another source of automatically-collected data is web server logs; these are vast collections of data relating the accessing of specific web pages (Hanna, 2004). Online learning researchers (Klassen & Smith, 2004; Lopes & David, 2006; Monk, 2005; Zaiane, 2001; Zorrilla, Menasalvas, Marin, Mora, & Segovia, 2005) have employed data mining techniques to gain useful insight from these data. Though, the

Table 1
Alternate sources of e-learning data

Data source	Method of analysis	Applied to community?
Text communication records: rich, high-level data; time intensive coding		
Lowes et al. (2007)	DM, SNA	Yes
Dringus and Ellis (2005)	DM	Yes
Mochizuki et al. (2005)	Real-time visualization, keyword recognition	Yes
Server log files: low-level data, high noise, difficult to organize		
Lopes and David (2006)	OLAP	No
Monk (2005)	Basic statistical	No
Zorrilla et al. (2005)	DM, OLAP	No
Klassen and Smith (2004)	Spreadsheet	No
Zaiane (2001)	DM	No
LMS log files: high-level data, more organized but still needs sorting		
Not real time		
Lowes et al. (2007, found of little use)	Basic statistical	Yes
Nurmela et al. (1999) (CSCL system log files)	SNA	Yes
Reffay and Chanier (2002)	SNA	Yes
Shen et al. (2007).	SNA	Yes
Silva and Vieira (2002) (platform-agnostic)	DM	Somewhat
Real time		
Moodie and Kunz (2003, proposed iLMS)	AI	Yes
Santos, Rodríguez, Gaudioso, and Boticario (2003, proposed CSCL system)	AI	Yes
Kosba, 'TADV' iLMS (2004)	AI	Somewhat
Mazza 'CourseVis' LMS tool (2004)	Visualization	Yes
Ueno 'Samurai' iLMS (2004)	DM, AI	Yes
Mazza and Milani, 'GISMO' Moodle module (2005)	Visualization	Yes

DM=data mining SNA=social network analysis AI=artificial intelligence.
OLAP=Online Analytical Processing (an analytic method similar to data mining).