Discovering Learners Behaviour Patterns From Log Files Using LSA

Iness Nedji Milat, Badji Mokhtar University, Annaba, Algeria
Hassina Seridi, LABGED Laboratory, Badji Mokhtar University, Annaba, Algeria
Abdelkader Moudjari, MISC Laboratory, Abdelhamid Mehri University, Constantine, Algeria

ABSTRACT

Recently, discovering learner behaviour has taken more attention in the field of e-learning. It aims to gain useful insights into the learning process of students despite the absence of direct interaction with teachers. In fact, the only available source of information in such environments is the log file that represents all possible interactions of learners with the e-learning system. This log file is characterised by the presence of noise, incomplete information, and a huge amount of data. In this article, a new approach based on learner trails analysis from the log file is proposed. It aims to discover the patterns of the real behaviour of learners and to determine their pedagogic orientations. The latent semantic analysis (LSA) method is used to extract the relationship between learners who have the same behaviour and to overcome the noise problem. The proposed approach has been validated using synthetic and genuine log files. The obtained results show the efficiency of the proposed method of discovering the behaviours of learners.

KEYWORDS

EDM, Learner Action, Learner Behaviour Pattern, Log File, LSA, Pedagogic Orientation, Trails Analysis

INTRODUCTION

In the context of e-learning systems, instructors are not confronted with their learners to supervise them. They do not have enough information about what are learners doing, and whether they have problems during the learning process or not. Consequently, instructors cannot take the appropriate decisions in time to help learners (Graf & Liu, 2010; Martínez et al., 2015; Pardo, 2014).

To deal with these issues, several works have been proposed to analyse learner behaviours. This analysis aims to acquire knowledge about the learning process and to understand the learner behaviours in the e-learning system. This knowledge can be used by different stakeholders such as learners, instructors, designers and pedagogues. For example, learners could get self-awareness about their learning progress. The Instructors could supervise their learners. Additionally, course designers and/or pedagogues can adapt and personalise the learning process (Bourguignon, 2007; Romero et al., 2013; Pardo, 2014; Mahajan, 2014; Martínez et al., 2015; Cereso et al., 2016).

Actually, the analysis of learner interactions serves as a mechanism to understand their behaviours in the e-learning system. This analysis is achieved by using log files-based processing (Bousbia et
al., 2010). Log files save all events that can occur during the learning process. However, these events could not be exploited directly by different stakeholders because they are low-level, noisy, incomplete, and huge data. For this, it is difficult to extract knowledge directly from these log files.

Among the techniques used in the field of e-learning, to extract useful knowledge from log files, the authors find Educational Data Mining (EDM) (Neuhauser, 2002; Romero & Ventura, 2010; Siemens & Long, 2011; Agosti et al., 2012; Greller & Drachsler, 2012; Stater et al., 2016; Bakhshinategh et al., 2018). This technique is used to well-understand learner interactions and to discover their behaviours (Cereso et al., 2016). EDM focuses on developing and applying computerised methods to detect patterns in large collections of educational data, which would otherwise be difficult or even impossible to analyse (Romero & Ventura, 2010; Peña-Ayala, 2013). These methods include data mining, machine learning, statistics, data visualization, and modelling methods (Baker & Siemens, 2014; Bakhshinategh et al., 2018).

In the last decades, several works based on EDM methods have been proposed to pattern student behaviour in Learning Management Systems (LMS). Most of these works focused on computing indicator variables such as participating rate, the effort provided during learning, self-assessment test, reading time spent, working time spent, discussion forum and/or final score, etc. (Hang & Zhang, 2008; Bovo et al., 2013; Kim et al., 2014; Cereso et al., 2016). However, the main drawback of using these indicators is that they do not reflect the real behaviour of learners in the e-learning system. For example, the reading time spent does not necessarily indicate the real time spent in reading a course material. Also, these variables would not be resistant to noise if the server is overloaded. These works, based on indicator variables analysis, are oriented for identifying and predicting specific goals, especially, learning styles, student achievement, dropout and at-risk, without describing learner behaviours.

The other works tackling the same problem are based on web usage mining techniques (Beauvisage, 2004; Bourguignon, 2007; Bousbia et al., 2010). The procedure is achieved by analysing the learning trails (paths) that provide knowledge about the activities performed by learners. These works are oriented to describe the process of learning; specifically the navigation behaviour.

In the present paper, authors have considered the learning trails to discover the patterns which describe learner behaviours. These trails represent the sequence of actions performed by learners during their learning process. The authors note here, that the trails could be pre-established as learning scenario or no. To automate the discovery task, they model the trail of learner behaviour with a graph in which the nodes represent learner actions and the arcs represent action transitions. In addition, EDM methods are applied to analyse learner interactions with the system regarding only the trails of learners without computing the indicator variables. The proposed approach aims to discover different patterns of learner behaviours and to determine their pedagogic orientations in order to well-understand the learning process.

On the other hand, the proposed approach uses a mathematical method to extract relationships between learners and their actions using only log files generated from the Moodle platform. This method is the Latent Semantic Analysis (LSA). The LSA is a process inspired by natural languages that consider the words which appear in the same context as being semantically related (Deerwester et al., 1990). It was initially developed as a tool for retrieving documentary information (Dumais, 1997). But very quickly, because of its good performances, its use has been extended to other areas such as knowledge acquisition (Redington et al., 1998; Lemaire, 1999; Dessus, 2000; Dumais, 2005) and recently knowledge discovery. In the latter area, Moudjari et al. (2014) used the LSA method for business protocol discovery which aimed at discovering the dynamic part (behaviour) of a web service. Slomovitz (2017) also used the LSA to discover hidden relations between Syslog messages, which can be used to automate the correlation process and help in the network event detection.

The key idea of the present work is to exploit only the information contained in the log files; not only for one learning session or specific learner, but the information of many sessions for many learners. This is made possible by using an LSA-based method. The information about learner
Related Content

Religion and Online Learning
www.igi-global.com/chapter/religion-online-learning/11983?camid=4v1a

Peer Coaching and Reflective Practice in Authentic Business Contexts: A Strategy to Enhance Competency in Post-Graduate Business Student
www.igi-global.com/chapter/peer-coaching-reflective-practice-authentic/27603?camid=4v1a

A Strategy to Expand the University Education Paradigm: Selling Online Class Resources
Richard Ryan (2004). Distance Learning and University Effectiveness: Changing Educational Paradigms for Online Learning  (pp. 235-252).
www.igi-global.com/chapter/strategy-expand-university-education-paradigm/8571?camid=4v1a
Problematic but Possible: Online Teaching and Learning in Post-registration Programmes
Maria Cassar and Josef Trapani (2014). *International Journal of Distance Education Technologies* (pp. 14-21).
[www.igi-global.com/article/problematic-but-possible/117179?camid=4v1a](www.igi-global.com/article/problematic-but-possible/117179?camid=4v1a)