

Behavior Usage Model to Manage the Best Practice of e-Learning

Songsakda Chayanukro¹, Massudi Mahmuddin², Husniza binti Husni³

¹Universiti Utara Malaysia, Malaysia, songsakda@gmail.com

²Universiti Utara Malaysia, Malaysia, ady@uum.edu.my

³Universiti Utara Malaysia, Malaysia, husniza@uum.edu.my

ABSTRACT

This study aims to find e-Learning users' behavior model that use data mining techniques to predict the successful learning behavior in utilizing e-Learning systems and to develop appropriate e-Learning users' behavior models that could be used broadly in other higher institutions. Due to the lack of suitable e-Learning user's behavior model for open source e-Learning system (Moodle) that could not be able to make a prediction for learning outcomes or performances. In this case, it is not useful enough for improving learners' performance which may cause failure in learning. Therefore, this research is conducted upon three main phases, which are data preparation, data extraction and model verification for generating a verification pattern. This pattern could be used as a direction for creating a more appropriate e-Learning users' behavior model.

Keywords: e-Learning, user behavior, data mining

I INTRODUCTION

In many academic institutions as well as commercial organizations, the system that can support and improve learning, such as institutions or universities nowadays, keep on developing or buying system for this purpose. Especially, Learning Management System (LMS) plays more crucial roles in distance learning because of its manageability. This system is an able to manage web log of the registered users, manage course catalogs, record data from learners and equipped with reports for management. For the distance learning education, LMS software is very economically and practicable. Beside that, this software can be used in many different phrases that can support users in terms of performing content preparation by keeping records of the users' history. An advantage of LMS open source software is easily adapting database structure especially the structure of a usage history so called web log. The web log is a hidden useful part, which is a helpful factor for developing a stable and appropriate e-Learning users' behavior model using data mining technique. To strengthen the e-learning system, Moodle (2011) continuously tries its best to make improvement of debug problem and implementation of more various new functions in order to fulfill the new version. Referred to www.moodle.org website

forum, there are several works that attempt to improve and develop novelty in various features for its new version which can be categorized into two aspects as e-Learning tools and e-Learning users' behavior models (Figure 1.)

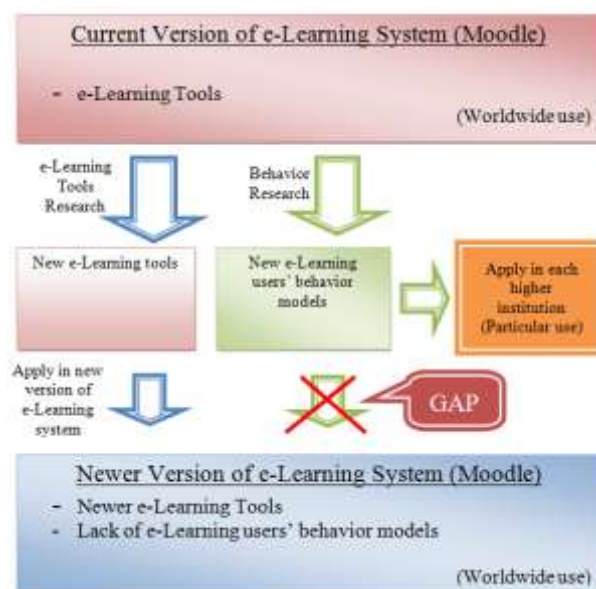


Figure 1. Gap of e-Learning Users' Behavior Model Development

As depicted in Figure 1., e-Learning tools are the functions to enhance on-line learning effectiveness, which synthesized based on Moodle staff e-Learning development and other researchers. Later on, this model can be downloaded and also can be set up for widely usage. Therefore, the users' behavior model could be improved positively which can effect to learners' higher performance. However, those models could not be applied to other higher institutions due to lack of readiness for developing their own suitable models for enhancing learners' performances broadly. Furthermore, the developed e-Learning users' behavior models still not utilize to the newer version of e-Learning system tools. Some higher institution researchers (Lingyan, Jian, Lulu, & Pengkun, 2010; Ribeiro & Cardoso, 2008) use their web logs for developing appropriate and efficiency model only for particular uses. In the same time, the study of e-Learning user's behavior evaluation that can enabled the new model which related between learning behavior and effect, and can be used to evaluate students in their e-Learning system (Lingyan, et al., 2010).

According to the researches on e-Learning users' behavior models (Chien-Ming, Chao-Yi, Te-Yi, Bin-Shyan, & Tsong-Wuu, 2007; Chun-Xia, Hui-Bao, Chang-Yi, & Yue-Xing, 2010; Ribeiro & Cardoso, 2008), Mentioned that, some study have been approached on the models which is very helpful for either teachers or learners to realize learning status and learning outcomes for learners' higher achievement. Nevertheless, the studies of the similarities of e-Learning users' behavior models from different higher institutions are deficient.

Therefore, the study of making understanding of the relationship between two different models could be approach to their similar pattern that will be useful for gaining the knowledge to develop an appropriate universal model.

II E-LEARNING

e-Learning has become the important part of learning system. Currently, there is an increasing interest in data mining and educational systems, making educational data mining as a new growing research community. The popularity of e-Learning has grown rapidly over the last decade in higher education (Dai & Zhang, 2008). e-Learning system can allow students to learn the lectures material, and experience the learning process through the network (Min, 2005). At the present, e-Learning is remarkably developed in order to produce effective learning outcome. This advanced system is not only initiated its tools to meet the learning activities in reality. Furthermore, it is also originated to create virtual classroom management systems for instance user authentication and classroom communication in order to foster an efficient virtual classroom.

Learning Management System (LMS) is system software for learning that enables the display of theoretical content in an organized and controlled way. It mainly consist of administration, content packing, synchronous and asynchronous communication tools, knowledge evaluation, tracking users (Sancristobal et al., 2010). LMS provides a platform to allow interactions between students and tutors, as well as among the peers. Most of the conventional pedagogic activities can be performed on e-learning environment (Hsien-Tang, Chih-Hua, Chia-Feng, & Shyan-Ming, 2009). In the same time, higher institutions look for the best LMS either open source that best suit in commerce. However, mostly LMS are developing to meet the standard pattern that can use with other such as Sharable Content Object Reference Model (SCORM) (scorm.com, 2012) that is a collection of standards and specifications for web-based e-Learning and it is the famous standard pattern.

III OPEN SOURCE E-LEARNING

Nowadays, the number of open source e-Learning system become increasing such as Moodle is one of the best LMS because of the designed based on social constructionist pedagogy. It has been widely adopted in 200 countries, has more than 40,000 registered sites, and the number of courses is in excess of 2,400,000 (Hsien-Tang, et al., 2009). Modular Object-Oriented Dynamic Learning Environment (Moodle) (Moodle.org, 2011) is a Course Management System (CMS). This is a free web-based application that educators can use to create effective online learning sites. This open source LMS contains Sharable Content Object Reference Model (SCORM) (scorm.com, 2012) standard ready (Ruiz Reyes et al., 2009). Thereby, the study on open source e-Learning user's behavior could be the broadly advantage. Table 1 presents mainly activities were appear in mostly open source e-Learning system that could use for e-Learning users' behavior studying.

Table 1. Activities of open source e-Learning group by subcategories (Graf & List, 2005).

Subcategories	Activities
Communication tools	Forum, Chat, Mail/Messages, Announcements, Conferences, Collaboration, Synchronous & asynchronous tools
Learning objects	Tests, Learning material, Exercises, Other creatable LOs, Importable LOs
Management of user data	Tracking, Statistics, Identification of online users, Personal user profile
Usability	User-friendliness, Support, Documentation, Assistance
Adaptation	Adaptability, Personalization, Extensibility, Adaptively
Technical aspects	Standards, System requirements, Security, Scalability
Administration	User management, Authorization management, Installation of the platform
Course management	Administration of courses, Assessment of tests, Organization of course objects

According to the study of Ribeiro & Cardoso (2008), the data obtained from the web log in Moodle LMS is not only important as a navigational framework but also provide relevant input for selective model construction which is crucial for tracking students' behavior. Their work shows the model, which is able to successfully predict students' final outcome while bringing useful feedback during course making.

IV LEARNING BEHAVIOR

Kebin, Feimin, Ming, Feng, & Xiaoshuang (2008) define e-Learning behavior as the long-distance independent learning behavior that take place in the learning environment which was constructed by information technologies. The learning portfolio is the e-Learning user's behavior data that provide the students with a specific method to evaluate their own learning situations. They include all the records of the students' activities during the learning process, such as their interaction with others, assignments, test papers, personal work collections, their discussion content, and online learning records (Chien-Ming, et al., 2007). The structure of the database from open source e-Learning system has illustrated in Table 2.

In order to make an e-Learning user behavior is more explicit, the developed e-Learning users' behavior model should be generic for generalization. Some researches design a kind of active e-Learning system was based on students' requirements and proposed the workflow of the system and design the function modules of the active e-Learning system. The designs focus on students activeness and include most active learning function or tools of e-Learning (Chun-Xia, et al., 2010). For this reason, it is crucial to discuss about the actual e-Learning users' behaviors in order to generate the useful model that can be able to generalize in e-Learning development.

V USER BEHAVIOR BY DATA MINING TECHNIQUE

There are many important data tables for e-Learning usage behavior study such as learner information, learner's action log and learner studying activities as presents in Table 2 that contained the necessary data for data mining process. A collection of normal web log could help to explain the phenomenon of e-Learning users' behavior comprehensively. In addition, it explains the e-Learning users' behavior in different periods of time and different user groups that suitable for many data mining technique algorithms.

Table 2. Important Moodle tables for doing data mining (Romero, Ventura, & Garcia, 2008).

Name	Description
mdl_user	Information about all the users.
mdl_user_students	Information about all students.
mdl_log	Logs every user's action.
mdl_assignment	Information about each assignment.
mdl_assignment_submissions	Information about assignments submitted.
mdl_chat	Information about all chat rooms.
mdl_chat_users	Keeps track of which users are in which chat rooms.
mdl_choice	Information about all the choices.
mdl_glossary	Information about all glossaries.

mdl_survey	Information about all surveys.
mdl_wiki	Information about all wikies.
mdl_forum	Information about all forums.
mdl_forum_posts	Stores all posts to the forums.
mdl_forum_discussions	Stores all forums' discussions.
mdl_message	Stores all the current messages.
mdl_message_reads	Stores all the read messages.
mdl_quiz	Information about all quizzes.
mdl_quiz_attempts	Stores various attempts at a quiz.
mdl_quiz_grades	Stores the final quiz grade.

VI USER BEHAVIOR MODEL DEVELOPMENT

The data derived from this study concerned with e-Learning users' behavior that could be affected to the model analyzing from the data mining technique more stable and more generalized. Therefore, the study of this information shall be used to develop more stable and more effective models of e-Learning users' behavior. In order to make a prediction model, web log is utilized for creating a useful model to determine e-Learning users' behavior.

Figure 2 shows a process of proposed approach consists are three steps of

- i) Data preparation process where the data sources is derived from.
- ii) Data extraction process that uses data mining techniques to find out the best models.
- iii) Model verification process that verifies between the output models from step two and the other model from another e-Learning system.

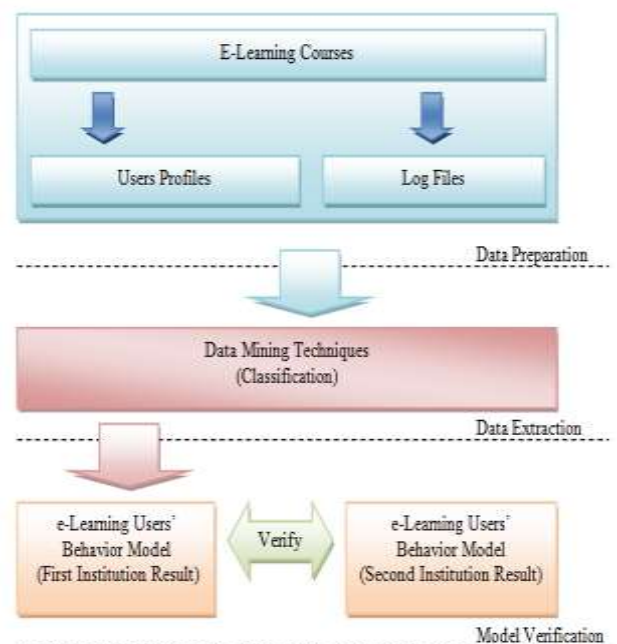


Figure 2. Proposed Methodology

A. Evaluation of e-Learning Usage Behavior

For discussing on the evaluation of the e-Learning users' behavior, there are several attracted researches proposed. The study of behavior factor evaluation was display the attributes in web log. Three groups of attributes (count, time, score) are the dimensions for user's behavior evaluation that could be classifying all web log attributes in these groups for data analysis process. It could be classified the attributes processing in three group as shown in Table 3.

Table 3. Group of Web Log's Attributes Processing (Lingyan, et al., 2010)

Group of Attributes Processing	Attributes' Activities
Count	<ul style="list-style-type: none"> ▪ The number of learning resource (TotalCount) ▪ The number of asking questions (QuesCount) ▪ The number of answering questions (AnsCount) ▪ The number of sending posts (bbsSentCount) ▪ The number of replying post (bbsAnsCount) ▪ The number of tests having been done (TestCount) ▪ The number of assignments (HomeworkCount)
Time	<ul style="list-style-type: none"> ▪ The average time of learning resource (TotalTime)
Score	<ul style="list-style-type: none"> ▪ The average score of the tests has been done (Test Score), which is divided into five levels: 'A' represents the score greater than or equal to 90 points, 'B' represents the score between 80 and 89 points, 'C' represents the score between 70 and 79 points, 'D' represents the score between 60 and 69 points, 'E' represents the score smaller than 60. ▪ The average score of assignment (HomeworkScore).

Table 3 presents the learner activities (web log's records) processing types that processes by activities counting, activities timing (period) and activities score. Thereby, these three groups of processing could be using in e-Learning users' behavior web log processing.

B. Data Mining on e-Learning Data

In order to extract the information from the huge database called web log, data mining technique plays a crucial role. As mentioned in the study of Wen-Hai (2010), the study of using data mining technique to excavate the client behavior pattern from web log files emphasis on analyzing client behavior pattern recognition system and its application for aiming at obtaining client information conveniently and automatically.

The support vector machine (SVM) is a supervised learning method that generates input-output mapping functions from a set of labeled training data. The mapping function can be either a classification function, i.e., the category of the input

data, or a regression function. For classification, nonlinear kernel functions are often used to transform input data to a high-dimensional feature space in which the input data become more separable compared to the original input space. Maximum-margin hyperplanes are then created. The model thus produced depends on only a subset of the training data near the class boundaries (Lipo, 2005). Thereby, the web log processing as show in Table 3 and using support vector machine (SVM) for evaluation is also the method that could be use in this study.

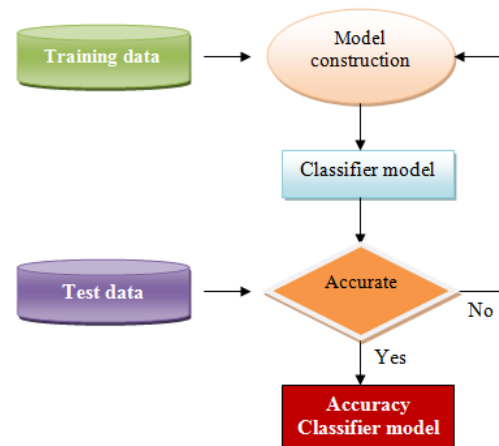


Figure 3. Accuracy classifier model finding

Step of data analysis and evaluation could be used from data mining processing step. After that, e-Learning system web log and user profiles from system databases are collected as a data for the study. Finally, the data will be proceeding by the following steps (Figure 3.)

The first step is data preparation. This step will be processed with three tasks as data cleaning, data selection and data transformation. In data transformation, will classify all attributes in the web log into three groups as activities counting, activities timing (period) and activities score (see in Table 3.).

The second step is data extraction. This step will be processed with classification technique. Support Vector Machine (SVM) is the one of classification technique that suitable for learner behavior analysis (Ribeiro & Cardoso, 2008) in order to develop learner behavior prediction model. This model will be used for predicting learners' grading results that can be effect to learners' performance improvement. This process will divide data obtained from the first step into two groups. Then a model will be constructed by SVM technique with training and test data up until an accuracy classification model ensures it.

The standard evaluation technique in situations where only limited data is available is stratified 10-folds cross-validation (Witten & Frake, 2005). Thus, the partition of data for training and test set will be use 10-fold cross-validation method, which has become the standard method in practical terms. Tests have also shown that the use of stratification improves results slightly. However, the partition number of in this study can be increasing up to the accurate error estimate, which target of predicting performance will determine the success rate at least 75% (Witten & Frake, 2005). All data analysis is shown in Figure 3.

C. Generalise e-Learning Usage Behavior Model

According to the problem statement of this study, there is an unknown similar pattern from difference user's behavior models that need to examine by following flow chart as illustrated in Figure 4.

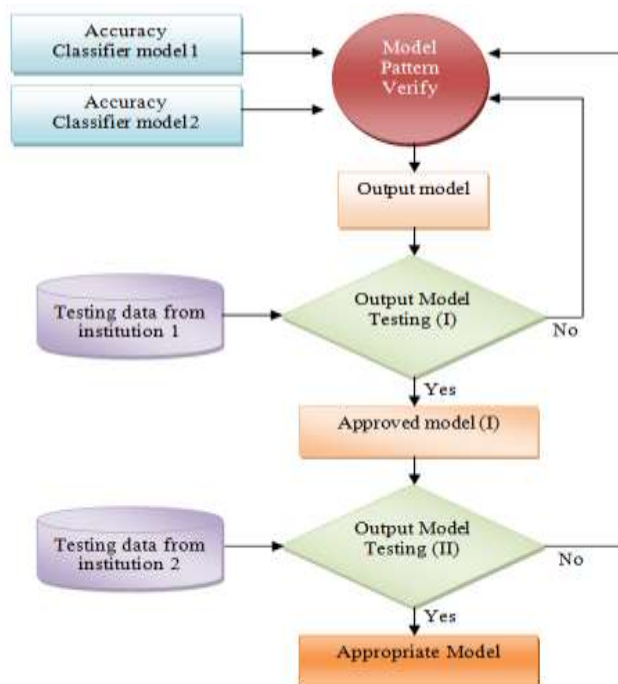


Figure 4. Model pattern verification

As depicted in Figure 4., according to the data extraction process, the results from both institutions' accuracy classification model will be proceeding for verifying model pattern. Consequently, the new output model will come out from this step.

The model pattern verification process is for adjusting the parameters and initial conditions of a model with another one in order to calibrate the new validity output model. Thereby, the output of this process is a pattern recording of the two models calibration.

The approved model (I) and appropriate model as shown in figure 4. are the outputs of cross-validation process, which target of predicting performance will determine the success rate at least 75% (Witten & Frake, 2005).

In order to test an accuracy of the models, test data from institution 1 is availed to find out whether it can be approved or not. Once it is not approved as a right model, the model patterns verify process will be taken simultaneously. On the other hands, if it is approved, it will be employed for the second accuracy checking.

From the approved model (I), test data from institution 2 is used to find out whether it can be approved or not. Concordantly, if it is not approved as a right model, the model patterns verify process will be taken again. Once it is approved, appropriate model will finally be discovered.

VII CONCLUSION

The main goal of this research is to develop appropriate models to describe the e-Learning users' behavior in order to meet an effectiveness of educational development. Furthermore, an approaching to the unknown similar pattern of e-Learning usage behavior model from different e-Learning systems is the important starting point to invent more appropriate general e-Learning usage behavior model. From the web log, users' activities history in e-Learning system is hidden in the most important factors. This study suggests new model that could explain the relevant of broader e-Learning users' behavior. At the same time, this study presents the useful and universal model that is hidden in the system history, which can be contributed for other higher institutions in term of using e-Learning system. Moreover, it is an advantage for the others higher institutions to use this model rather than create a new model which is incur time and cost for developing the new model of e-Learning users' behavior as well as the pattern of two models verifying could be the base for further study in users' behavior function of newer Moodle version.

REFERENCES

- Chien-Ming, C., Chao-Yi, L., Te-Yi, C., Bin-Shyan, J., & Tsong-Wuu, L. (2007, 10-13 Oct. 2007). *Diagnosis of Students' Online Learning Portfolios*. Paper presented at the Frontiers In Education Conference - Global Engineering: Knowledge Without Borders, Opportunities Without Passports, 2007. FIE '07. 37th Annual.

- Chun-Xia, Q., Hui-Bao, C., Chang-Yi, L., & Yue-Xing, S. (2010, 22-24 June 2010). *Design an active e-learning system*. Paper presented at the 2nd International Conference on Education Technology and Computer (ICETC), 2010.
- Dai, S., & Zhang, P. (2008, 16-18 April 2008). *A data mining algorithm in distance learning*. Paper presented at the CSCWD 2008. 12th International Conference on Computer Supported Cooperative Work in Design, 2008.
- Graf, S., & List, B. (2005, 5-8 July 2005). *An evaluation of open source e-learning platforms stressing adaptation issues*. Paper presented at the ICALT 2005. Fifth IEEE International Conference on Advanced Learning Technologies, 2005.
- Hsien-Tang, L., Chih-Hua, W., Chia-Feng, L., & Shyan-Ming, Y. (2009, 13-15 Nov. 2009). *Annotating Learning Materials on Moodle LMS*. Paper presented at the ICCTD '09. International Conference on Computer Technology and Development, 2009.
- Kebin, H., Feimin, L., Ming, Z., Feng, W., & Xiaoshuang, X. (2008, 25-26 Sept. 2008). *Design and Implement on E-learning Behavior Mine System*. Paper presented at the WGEC '08. Second International Conference on Genetic and Evolutionary Computing, 2008.
- Lingyan, W., Jian, L., Lulu, D., & Pengkun, L. (2010, 23-25 July 2010). *E-Learning Evaluation System Based on Data Mining*. Paper presented at the 2nd International Symposium on Information Engineering and Electronic Commerce (IEEC), 2010
- Lipo, W. (2005). *Support Vector Machines: Theory and Applications* (Vol. 177).
- Min, J. (2005, 19-21 May 2005). *Development of an e-learning system for teaching machining technology*. Paper presented at the Proceedings of the 2005 International Conference on Active Media Technology, 2005. (AMT 2005).
- Moodle.org. (2011). About Moodle Retrieved 26/10/2011, 2011, from [http://docs.moodle.org/20/en/About Moodle](http://docs.moodle.org/20/en/About_Moodle)
- Ribeiro, B., & Cardoso, A. (2008, 12-15 Oct. 2008). *Evaluation system for e-learning with pattern mining tools*. Paper presented at the SMC 2008. IEEE International Conference on Systems, Man and Cybernetics, 2008.
- Romero, C. b., Ventura, S., & Garcia, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368-384.
- Ruiz Reyes, N., Vera Candeas, P., Galan, S. G., Viciano, R., Canadas, F., & Reche, P. J. (2009, 22-24 June 2009). *Comparing open-source e-learning platforms from adaptivity point of view*. Paper presented at the EAEEIE Annual Conference, 2009.
- Sancristobal, E., Castro, M., Harward, J., Baley, P., DeLong, K., & Hardison, J. (2010, 14-16 April 2010). *Integration view of Web Labs and Learning Management Systems*. Paper presented at the Education Engineering (EDUCON), 2010 IEEE.
- scorm.com. (2012). Shareable Content Object Reference Model (SCORM) Retrieved 25/04/2012, from <http://scorm.com/scorm-explained/>
- Wen-Hai, G. (2010, 11-14 July 2010). *Research on client behavior pattern recognition system based on Web log mining*. Paper presented at the ICMLC 2010 International Conference on Machine Learning and Cybernetics
- Witten, I. H., & Frake, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (Second ed.). San Francisco, CA: Morgan Kaufmann.