The Use of Tin Can API for Web Usage Mining in E-learning Applications on the Social Network

W. M. Amir Fazamin W. Hamzah[1], Noraida Haji Ali[2]
Md. Yazid Mohd Saman[3] and Mohd Hafiz Yusoff[4]
School of Informatics & Applied Mathematics
Universiti Malaysia Terengganu
K. Terengganu, Malaysia
[1]amirfazamin@gmail.com, {[2]aida, [3]yazid,
[4]hafiz.yusoff}@umt.edu.my

Azliza Yacob
Faculty Computer, Media & Technology
TATi University
Terengganu, Malaysia
azliza@tatiuc.edu.my

Abstract— **Nowadays, the usefulness of the social network is very widespread. Besides used to socialize, business and advertising, the social network is also used for learning. In this study, Facebook is used as a medium of learning. E-learning applications have been embedded into Facebook as Facebook applications. This is to facilitate access to e-learning applications. Previous research has shown that web usage mining method could only track users who access e-learning applications on the social network. They did not track the behaviour of users in e-learning applications. Therefore, the use of Tin Can API is proposed to overcome this issue. Integrating e-learning applications with Tin Can API enables user behaviour can be tracked and recorded. In this paper, the use Tin Can API for web usage mining in e-learning applications on the social network is explored.**

Keywords - E-learning applications, web mining, web usage mining, Tin Can API

I. INTRODUCTION

The use of the social network (SN) has increased intensely in recent years. Social networking sites like Facebook are in the highest position achieved, followed by Twitter and LinkedIn [1]. Data show that users spend quite a lot of time when visiting Facebook. Teenager makes Facebook as their social networking tools. Social networking allows students to build profiles, build relationships, and discussion through the website. Learning through a SN can be a valuable part of an educational curriculum. As with all learning, students need guidance and opportunities to reflect on their work. A teacher can use SN as a supportive content for learning. Visualization of the concepts of learning subjects into SN will attract students' interest [2]. The development of e-learning applications on the SN must meet the requirements of students. It can be predicted with conducting studies of the behaviour of students [3]. One method to explore the behaviour of students in using e-learning applications is web mining technique.

Web mining is a data mining technique that is often used to analyse the students' behaviour in an e-learning environment [3]. Tracking the behaviour of students in using e-learning applications is important for knowing the behaviour of students. Analysis of students' behaviour in web usage mining will help the developers to modify, or make improvements of e-learning applications [4]. Web log file is the main sources that used for web usage mining. The behaviour of students when browsing the website can be recorded in the web log file [5-6]. However, when students access the e-learning applications that are embedded in SN websites, web log files only displays the records of students attempted and finished. The web log file did not explain the students' behaviour inside the e-learning applications on the SN.

Previously, a lot of method to track the students' experience that using e-learning has been developed. One such method is with the use of Tin Can API. The Tin Can API, or also known as the Experience API (xAPI) is an e-learning software specification developed by Advanced Distributed Learning (ADL) and Rustici Software [7]. Tin Can API allows all types of learning experience in using e-learning applications is tracked and recorded. The results of learning experiences are stored in the Learning Record Store (LRS), which is in the Learning Management System (LMS), or by itself. Based on the above facts, it is important to do research on methods to track the behaviour of students by using Tin Can API. Furthermore, studies involving web usage mining in e-learning applications on the SN did not yet widespread. Therefore, the objective of this paper is to explore the use of Tin Can API for web usage mining in e-learning applications on the SN.

II. RELATED WORK

A. E-learning Applications

E-learning is often related with areas that use technology to deliver learning programs and training. Typically used to describe media such as CD/DVD-ROM, internet/intranet, audio/video and mobile learning. The example of e-learning is the computer and network-enabled transfer of skills and knowledge. According to [8], e-learning should be defined basically as learning and teaching facilitated online through the technologies of the network. E-learning applications can be described as a typical a small program, the specialized program can be downloaded onto mobile devices or embedded program in website. The high-quality design of e-learning applications is helpful to promote the intrinsic motivation and learning quality [9]. In order to expand of e-learning, many researchers have developed e-learning applications for SN. E-learning applications on the SN usually in a format of Flash, or Html5 file embedded in SN sites.

978-1-4799-8969-0/15/$31.00 ©2015 IEEE

B. Web Mining

Data mining techniques to web or called as web mining can be defined as the discovery and analysis of useful information from the World Wide Web. It can be used for the analysis of e-learning applications on the SN [10]. Generally, web mining methods are divided into three types: web content mining, web structure mining, and web usage mining [11]. Web content mining can be described as a web mining technique to analyse the resources in the web and online databases. In web mining, the technique that can be used to analyse the links and structure of websites is web structure mining. Web usage mining can be described as a web mining technique that can be used to analyse of user access patterns or navigation behaviour of websites.

C. Web Usage Mining

Web usage mining is the discovery of significant patterns from data generated from client-server interaction on the web. The information about users' behaviours and their usage patterns knowable through web usage mining [12]. Analysing users' information from log file data can help improving the design and characteristics of a website. Mining information from web usage data also allows the restructuring of website management to be more systematic and interactive. There are three main sources for log file in web usage mining: web servers, proxy servers, and web clients [3][5].

- Web servers
  Log files on web servers typically contain basic information such as name, IP address, date and time of the request, and activity of the user.
- Proxy servers
  The proxy server takes requests from the user and passes them to the Web server. Then, the proxy server returns to the user and the results delivered to them by the web server.
- Web clients
  Web usage data can be tracked also on the client side by using special software that records web usage.

D. Web Usage Mining Process

The process of web usage mining can be divided into four different phases: data preparation, pre-processing, pattern discovering, and pattern analysis [12][13].

- Data preparation
  In this phase, the web log data will be cleaned, and filtered. Data that is not relevant and is not required will be removed. The information required will be identified.
- Pre-processing
  Pre-processing phase is the step to convert the information on usage, content and the structure of the information contained in various existing data sources into the data format required for pattern discovery [14]. In this phase, it is necessary to ensure user activities in the web server is readable and consistent.
- Pattern discovery
  Pattern discovery consists of different techniques from various fields such as machine learning, statistics, data mining and pattern recognition used to extract patterns of usage from web log data [15]. Pattern discovery can be done through statistical analysis techniques, association rules, sequential pattern analysis, classification, and clustering
- Pattern analysis
  This phase is to confirm and explain the patterns produced on pattern discovery phase [5]. The results of analysis can be represented in various ways, such as graphs, charts, forms, and tables.

E. Tin Can API

Tin Can API is a modernized version of SCORM (Sharable Content Object Reference Model). The purpose of Tin Can API is to store and provide access to learning data. It enables tracking of learning experiences, including conventional learning data, such as scores, or completion of the task [16]. It also recorded of learners' behaviour in using e-learning, like page navigation, reading notes or answered the quiz. It is a specification that describes the interface, and the rules of retrieval that developers can implement to track learning experience. This can be done by allowing statements of experience to be delivered to and stored in a Learning Record Store (LRS).

LRS is a central repository which stores all statements about learning experiences [17]. LRS enable the in-depth reporting and analysis on learning activities. Learner will access e-learning applications through a website that embedded e-learning applications. Then, the behaviour of learners using e-learning applications is recorded in the Tin Can statement. These statements are delivered to and stored in the LRS. The Tin Can statements format is based on activity types, and activity streams (Actor, Verb, and Object). The examples of activity types are module, course, cmi.interaction, and objective. Descriptions of the activity types are shown in the Table 1 [18]. Meanwhile, for the activity streams, the actor is the agent of the statement; that is, learner, an instructor, teacher, or group.

TABLE I. DESCRIPTIONS OF THE ACTIVITY TYPES

Activity types	Descriptions
Module	Describes a single slide/page within a course.
Cmi.interaction	A scored question.
Objective	Explaining the objectives of course. Each course can have one or more objectives.
Course	The e-learning applications

The verbs describe the action of the statement, such as: *attempted, experience, answered, completed, passed*, and *failed.* The object is what the Actor interacted with, a note, a quiz, or a class. Table 2 shows the descriptions of the Tin Can verbs [18].

TABLE II. DESCRIPTIONS OF THE TIN CAN VERBS

Verbs	Activity types	Description
Attempted	Course	Notifies the LRS that a course has begun.
Experienced	Module	Record a slide/page has been viewed.
Answered	Cmi.interaction	Represents that a question has been answered for a question.
Completed	Course, Objective	Used when a course or objective has been completed and completion is based on the number of slides/pages viewed.
Passed	Course, Objective	Notifies the LRS that a course or objective has been completed with a passing score. This verb is only used when achievement is based on a set of scored questions.
Failed	Course, Objective	Notifies the LRS that a course or objective has been completed with a failing score. This verb is only used when achievement is based on a set of scored questions.

III. THE PROCESS OF THE USE TIN CAN API FOR WEB USAGE MINING

In this section, the process of the use Tin Can API for web usage mining in e-learning applications on the SN were explained. The explanations of this process were divided into five steps, namely: design model, development of e-learning applications, Tin Can API integration, embedding of e-learning applications into the SN site, and implementation.

A. Design Model

Design model was a clear picture of the components involved in the use of the Tin Can API for web usage mining in e-learning applications on the SN. Figure 1 shows the design model. There were five components involved in the design model:

- Student or teacher who use e-learning applications through the SN.
- The SN site which embedded e-learning applications. The e-learning applications design consists of the menu and navigation, learning guides, learning contents, the high score table, and quiz.
- An LRS, where learning data were stored.
- The reporting tool was used to analyse the collected data. Normally, the results of the analysis will be visualized in graphical forms.
- The Learning Management System (LMS) also requires the learning data stored in the LRS. There was also an integrated LMS with LRS.

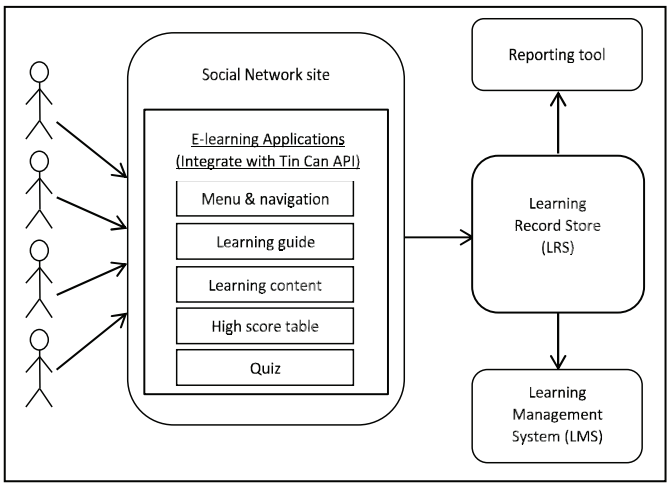

Figure 1. The design model

B. The Development of E-learning Apps

E-Learning applications for the SN was developed using multimedia authoring software Articulate Storyline. This software was a compliance software of Tin Can API. The use of compliance software of Tin Can API was important to enable the tracking of the user's experience with the contents of e-learning applications. E-learning applications that have been developed must be published for Tin Can API. This was because to enable e-learning applications can be integrated with the Tin Can API script. The contents of e-learning applications consist: menu and navigation, learning guide, learning content, the high score table, and quiz. Generated application format was Flash file or Html5. It was so that users can access a variety of applications with a modern web browser.

C. Tin Can API Integration

In this study, cloud hosting was used for e-learning applications. It was for easy access and flexible. E-learning applications were integrated with Tin Can API to track and record the users' experience. Figure 2 shows the configuration script of e-learning applications with Tin Can API. There were three compulsory elements that must be entered in the configuration file:

- Config.endpoint
  The LRS web address, where learning data was sent and stored.
- Config.authUser
  Id for the LRS admin or user.
- Config.Password
  Password for the LRS admin or user.

```
//globals: equal, responseText, statement, ok, deepEqual, QUnit, module, asyncTest, Util,
/*jslint bitwise: true, browser: true, plusplus: true, maxerr: 50, indent: 4 */
function Config() {
        "use strict";
}
Config.endpoint = "https://lrs.aplikasipembelajaran.com/data/xAPI/";
Config.authUser = "de35d79f09f2287d6e2020e31e3c75c306996fa9";
Config.authPassword = "cdd428dcb764e914d7330c6f0eaf3d4cb3875aeb";
Config.actor = {"mbox":[""], "name":[""] };
Config.registration = "2981c910-6445-11e4-9803-0800200c9a66";
```

Figure 2. Configuration of e-learning applications

All the required information was obtained from the LRS used. This configuration was important to launch e-learning applications and connects it with the LRS. The configuration files and Tin Can API scripts must be in the same hosted with e-learning applications. LRS used was Learning Locker, which was open-source software. LRS was hosted separately from e-learning applications.

D. Embedding of E-learning Apps into the SN Site

SN site used in this study was Facebook. E-learning applications were embedded into Facebook as Facebook applications. The student or teacher able to access e-learning applications when they were using SN site. They can explore the entire contents of e-learning applications with ease. With existing relationships among teacher and student through SN, the use of e-learning applications will bridge the relationship between teacher and student. This was because they can use the SN site for the discussion and sharing information about e-learning applications. Figure 3 shows the embedding of e-learning applications into the SN site.

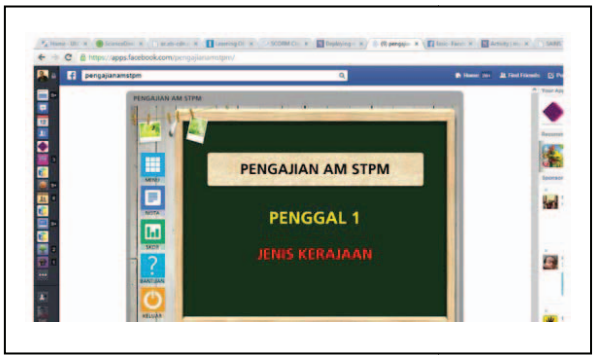

Figure 3. Embedding of e-learning applications into SN site

E. Implementation

Refers to the design model in Figure 1, the implementation of the use of Tin Can API for web usage mining in e-learning applications on the SN consists of four steps. There were:

- Student/teacher uses SN and accessed e-learning applications through SN.
- Tin Can API tracks all the learning data in the use of e-learning applications and sent record to LRS.
- Data was sent from LRS to a reporting tool for analysis. However, the LRS also can analyse data on its own.
- Data was sent from LRS to the LMS if there was demand.

Tin Can statements were used as the data source for web usage mining. The raw statements of Tin Can were shown in Figure 4.

Figure 4. Tin Can statements

IV. RESULT AND DISCUSSION

In this paper, the description of the process of data preparation and data pre-processing was not shown. It was because the study focuses on exploration of the use Tin Can API to get the data from the e-learning applications on the social network. However, the descriptions of the process of pattern discovery and pattern analysis were shown in this section. Data source was from the use of the e-learning applications of the *Pengajian Am* subject from a school in Malaysia. The period of data gathering for the use of e-learning apps on the SN was about one month. The pattern discovery process performed using statistical analysis and sequential pattern. The part of results from the analyses were described as below.

A. Statistical Analysis

Statistical analysis was used for extracting knowledge about users of e-learning applications. Typically, the extracted data for statistical analysis of usage, such as: *attempted, experience, answered, completed, passed,* and *failed*. The Tin Can statements were the type of qualitative data. Enumeration method was used to quantify the qualitative data for statistical analysis. The bar chart in Figure 5 shown the frequency of *attempted*. It has shown how often the students *attempted* e-learning applications. From the descriptive statistics (n=24), the mean of *attempted* was 11.3. The median was 8.5, and the mode was 3.0. *Attempted* had a standard deviation of 6.8.

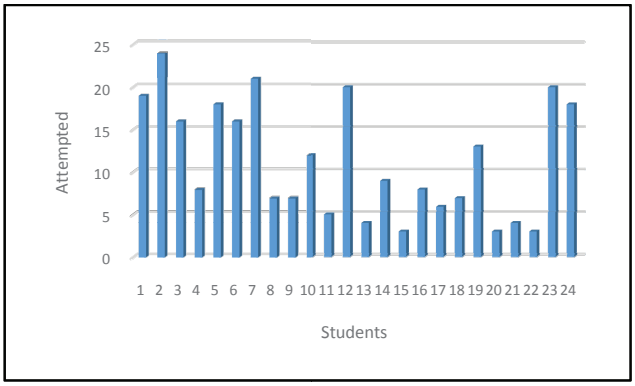

Figure 5. The frequency of *attempted*

The bar chart in Figure 6 shows the frequency of *experienced*. It describes the frequency of students to interact with components in e-learning applications such as: learning guides, learning contents, view high score, and quiz. From the bar chart, the highest frequency of *experienced* was the learning contents. Meanwhile, the lowest frequency of *experienced* was the learning guides.

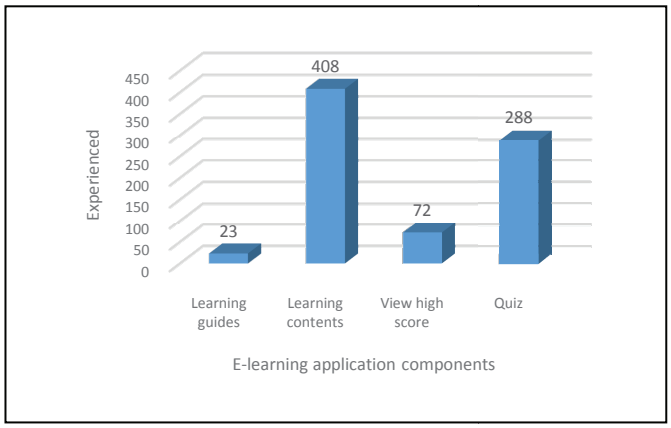

Figure 6. The frequency of *experienced*

B. Sequential Pattern

Sequential pattern was used to identify patterns of user's behaviour. The results of analysis can be used to predict the user's behaviour in the future. Individual behavioural learning patterns can be known when the detailed analysis conducted with the learning data. It describes how the behaviour of a student or individual who use e-learning application from start to finish.

Figure 7 shows the learning patterns of the student "A" who passing quiz at all the three attempts. In this study, the movement of a student in e-learning applications was referred as the steps. A student was *experienced* the components of learning guides such as syllabus and user guides when

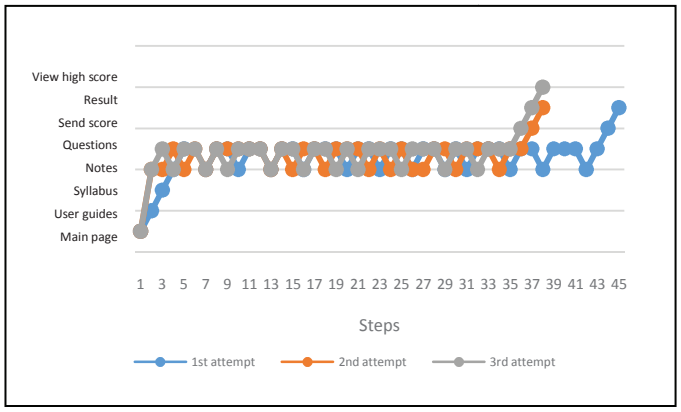

Figure 7. Learning patterns of the student "A"

using it for the first attempt. However, at the second and third attempt, he was not *experienced* learning guides. He was found refers to the notes when answering quiz questions in all attempts.

The learning patterns of the student "B" who fail the quiz at the first and second attempts was shown in Figure 8. However, he was passing a quiz at the third attempt. He rarely refers notes at the first and second attempt. Even so, at the third attempt, he always refers note when answering a quiz. He also was *experienced* learning guides at the third attempt.

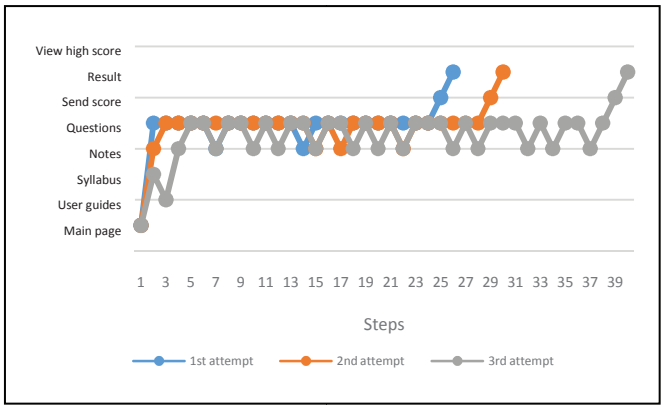

Figure 8. Learning patterns of the student "B"

Based on the learning patterns of student "A" and "B", the frequency of *experienced* the learning contents may influence to pass a quiz.

V. Conclusion

In this paper, the study of the use of Tin Can API for web usage mining in e-learning applications on the SN has been carried out. The method to track the behaviour of students who use e-learning applications that embedded in the SN site was explored. This was done by integrating e-learning applications with the Tin Can API. The process of implementation was explained in details. The use of Tin Can API for web usage mining enables the discovery of learning behaviour patterns with rapidly. This was because the resulting learning data were in the form of statements which easy to be analysed. The results of pattern analysis were useful for developers to improve e-learning applications through enhancements of design, learning content, and implementation.

In the future, more research work was required in the area of the use Tin Can API for mobile data mining. The integration of Tin Can API with mobile applications of e-learning enables big learning data can be accessed and analysed. In addition, the use of mobile applications was quickly growing.

References

[1] G. Fulgoni, "The Digital World in Focus - comScore, Inc," in The Digital World in Focus, 2013.

[2] M. K. Kabilan, N. Ahmad, and M. J. Z. Abidin, "Facebook: An online environment for learning of English in institutions of higher education?," Internet High. Educ., vol. 13, no. 4, pp. 179–187, Dec. 2010.

[3] W. Hu and C. Lee, “World Wide Web Usage Mining Systems and Technologies,” vol. 1, no. 4, pp. 53–59.

[4] B. Liu, Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, pp. 527–603,81–83, 2011.

[5] A. Upadhyay and B. Purswani, “Web Usage Mining has Pattern Discovery,” vol. 3, no. 2, pp. 1–4, 2013.

[6] A. Marengo, A. Pagano, and A. Barbone, “Data Mining Methods to Assess Student Behavior in Adaptive e-Learning Processes,” 2013 Fourth Int. Conf. e-Learning "Best Pract. Manag. Des. Dev. e-Courses Stand. Excell. Creat., pp. 303–309, May 2013.

[7] Á. Blanco, Á. Serrano, and M. Freire, “E-Learning Standards and Learning Analytics,” pp. 1255–1261, 2013.

[8] S. Hrastinski, “Asynchronous and Synchronous E-Learning,” no. 4, pp. 51–55, 2008.

[9] X. Cao, F. Wang, and Z. Zheng, “The Experimental Research on E-Learning Instructional Design Model Based on Cognitive Flexibility Theory,” Phys. Procedia, vol. 25, pp. 997–1005, 2012.

[10] S. Chakrabarti, “Mining the Web: Discovering Knowledge from Hypertext Data,” Online Information Review, vol. 27. pp. 291–291, 2003.

[11] R. Cooley, B. Mobasher, and J. Srivastava, “Web Mining: Information and Pattern Discovery on the World Wide Web,” in Ninth IEEE International Conference on Tools with Artificial Intelligence, pp. 558–567, 1997.

[12] B. Han, X. Hao, and C. Liu, “The design and implementation of user behavior mining in E-learning system,” Int. Conf. Autom. Control Artif. Intell. (ACAI 2012), pp. 2078–2081, 2012.

[13] N. Singh, A. Jain, and R. S. Raw, “Comparison analysis of w eb u sage m ining,” vol. 3, no. 4, pp. 137–147, 2013.

[14] J. Srivastava, R. Cooley, M. Deshpande, and P.-N. Tan, “Web usage mining: Discovery and applications of usage patterns from web data,” ACM SIGKDD Explor. Newsl., vol. 1, pp. 12–23, 2000.

[15] N. Zhong, Y. Li, and S. T. Wu, “Effective pattern discovery for text mining,” IEEE Trans. Knowl. Data Eng., vol. 24, pp. 30–44, 2012.

[16] M. Aberdour and A. Downes, “TIN CAN do for me?,” April, pp. 53–54, 2013.

[17] Kristy Murray, Peter Berking, Jason Haag, and Nikolaus Hruska, “Mobile Learning and ADL’s Experience API, Connections: The Quarterly Journal 12:1, 45-49, 2012.

[18] Articulate Storyline. (2014, Jun 11). Implementing Tin Can to support Articulate content [Online]. Available: https://en-uk.articulate.com/tincanapi/