Threshold value in automatic learning style detection

Norazlina Ahmad\textsuperscript{a}, Zaidatun Tasir\textsuperscript{b,*}

\textsuperscript{a}Department of Polytechnic Education, Persiaran Perdana Presint 4, 62100, Putrajaya, Malaysia
\textsuperscript{b}Faculty of Education, Universiti Teknologi Malaysia, 81310 Skudai, Malaysia

Abstract

Previous studies demonstrate that automatic approach is a better approach to identify learning styles in online learning because it is based on the actual students' behavior pattern while learning. However, the threshold value for every behavior pattern has to be obtained to identify students' learning style. The purpose of this study is to determine the threshold value based on patterns of behavior for 20 students who studied the Interactive Multimedia course. Patterns of behavior were analyzed based on students' online learning activities log file. The findings indicate that the threshold values are different depending on for the course structure, subject and experience of different students. Therefore, the determination of the threshold needed to ensure the measurement dimensions of online learning is based on actual usage of the investigated features of the course.

Keywords : Learning styles; automatic detection; literature-based approach

1. Introduction

The advancement of technology in the information system gives a profound impact on education, especially in online learning [1, 2]. A lot of effort have been done to ensure optimum online learning, i.e. the use of adaptive approach in supporting individual differences and needs [3, 4]. One of the parameters of individual differences used in adaptive online learning is based on learning styles. Most educational theorists and researchers believe that learning styles play an important role in the learning process in which it could make the learning easier and more effective [5-7]. By identifying students’ learning styles, teachers could design classroom activities and provide learning materials that are in line with the students’ learning styles [8]. In addition, students would be able to recognize their own learning strategies when they know the strengths and weaknesses of their learning styles [9].

James and Blank (1993), define learning style as ‘the complex manner in which, and the conditions under which, learners most efficiently and most effectively perceive, process, store and recall what they are attempting to learn’. There are many models of learning style refers to the variety and interest in the field of model designers. However there are five models of learning styles that are often used by engineering students in the previous studies: Myers-Briggs Type Indicator [10], Kolb [11], Felder and Silverman [8], Herrman [12] and Dunn and Dunn [13]. Nevertheless, the model that is used commonly is Felder and Silverman learning styles model [14, 15]. This model is able to analyze clearly the dimensions of learning styles based on a scale of range between +11 to -11 [16].

*Corresponding author. Tel.: +6-07-5534393; fax: +6-07-5534403.
E-mail address: p-zaid@utm.my
In addition, the questionnaire (Index of Learning Style, ILS) also has acceptable reliability and validity [17-19], as proposed by Kuljis and Liu [20], Felder and Silverman model of learning is suitable for online learning. Hence, the purpose of this study is to determine the threshold value based on the actual online learning behaviour pattern in order to identify learning style regarding to Felder and Silverman model.

2. Learning Style Model

Felder-Silverman learning style model was exclusively designed for engineering students [8]. This model was to identify students’ learning styles and help instructors to formulate approaches to teaching to be in line with a range of learning styles of students. There are five dimensions of learning style models: perception, input, organization, processing and understanding. However, there are only four dimensions found in the instrument used to determine learning styles [21]. According to the Felder-Silverman learning style, there are four dimensions of learning styles, namely: -(i) the perception or type of information that becomes a choice of someone while understanding a piece of information be it through sensory or intuitive, (ii) the most effective input or sensory channels that receive either visual or auditory perception, (iii) the active or reflective method of processing information and (iv) the sequence or global way a person understands a knowledge.

Richard M. Felder and Barbara A. Solomon created a Learning Style Index (ILS) in 1991 based on the Felder-Silverman learning style model [16]. This ILS questionnaire consists of 44 items and was placed on the World Wide Web in 1997 [21]. The instrument was designed with only four dimensions of learning style models: active / reflective, sensing / intuitive, visual / verbal and sequence / global. Each dimension in the questionnaire has 11 items. ILS scale is a bipolar scale with option ‘a’ brings value +1 and ‘b’ gives value -1. The total score for each dimension produces a scale from +11 to -11[16]. Scale value of 1 to 3 shows a balance between the two learning styles in one dimension, value of 5 to 7 refers to moderate learning styles while value 9 to 11 illustrates strong learning styles.

Although the Felder and Silverman model has four dimensions of learning styles, this study however, would only focus on the dimensions of information processing which are active and reflective. Processing dimension describes how to process the favourite information of a person whether it is active (actively involved in learning) or reflective (make observations on their own). Students with active learning styles are easy to understand information through learning activities that allow them to actively use the learning materials like doing experiments. Students with this type of learning style love to manipulate the learning materials, to try and test something new, give opinions and discuss in groups. However, the reflective learning style’s students are more comfortable making observations by investigating and manipulating learning materials or activities on their own. Next, to match learning materials and activities with appropriate teaching styles, learning styles should be identified. The next sections will describe various approaches that can be used to identify students’ learning styles.

3. Identifying Learning Style

According to Brusilovsky [22], there are two approaches that can be used to identify the styles of learning, which are collaborative and automatic. Collaborative approach is based on the questionnaire basis while automatic approach is based on the behaviour patterns during online learning. Collaborative approach is said to be less accurate because the students tend to be nonchalant when answering the questionnaires, which would affect the validity of the results [23]. Unlike collaborative approach, automatic approach is considered to be better in terms of the data accuracy since it is based on the actual behaviour of students [1, 24]. Yet, automatic approach has its own weaknesses. It takes a lot of time in getting the behaviour patterns of students who participate in the online learning and sometimes the patterns of the behaviour obtained from the data are not strong enough [25-27]. Accordingly, the study of learning styles by using automatic approach need to be done as it is believed to be able to identify learning styles more accurately. Besides, it could help the process of updating the students’ model more effectively in providing immediate actions based on the student needs.
In identifying learning styles via automatic approach, there are several methods used. Such methods differ in attributes (personality factors, behavior, etc.), analysis techniques (Bayesian networks, decision trees, etc.) and infrastructure (Learning Management System (LMS) or other system). Previous studies have shown that there are two methods used to identify learning styles in automatic approach: data-driven method and literature-based method [5, 26, 28]. Data-driven method aims at building a model that imitates the ILS questionnaire and uses sample data to construct a model [26]. While the literature-based method is to use the behaviour of students in order to get hints about their learning style preferences [5]. Neural networks [28], decision trees and Hidden Markov Model, fuzzy clustering [29] and Bayesian networks [26] are the techniques that based on the data-driven approach. The literature-based method on the other hand uses only simple rule-based to calculate learning style from the number of matching hints [5, 24, 27, 30].

The literature-based method which was proposed by Graf, Kinshuk and Liu [5] is said to be practical and widely used regardless the courses since the focus is more on the content of the objects such as quizzes, tests, assignments and forum which are the most featured items used in most open source LMS [1, 24, 31]. Besides that, the calculation of learning styles in this study is based on the simple rules and does not involve the design of the system. Therefore, it could be used in all learning situations which are LMS based [24]. Graf [32] conducted a comparative study of both data mining and literature-based methods and found that the literature-based method resulted in more accurate data of identifying learning styles compared to the data-driven method. However, the literature-based approach has its own weakness where there is a possibility that there might be other patterns that are not taken into account in the calculation which are also the behavioral indicators of learning styles [24, 33].

3.1. Literature Based Approach

Literature-based approach used the behavioral patterns found in students’ log data while using learning management system (LMS) to detect learning style. The technique used is matching the behavioral patterns of online learning with all the characteristics of learning style dimensions mentioned in the Felder and Silverman learning styles model. The simple rule-based method was used to calculate learning style based on the match number of behavior patterns [5].

In order to determine hints of the behavior pattern valued between 0 and 3, the relevant patterns of behavior while using LMS to be identified based on the characteristics of Felder and Silverman learning style model dimensions. In addition, the selection of the incorporated characteristics and patterns of behavior should also taken into consideration the probability of matching behavior patterns with high LMS characteristics used by the teacher. Based on a study done by Graf, Kinshuk and Liu [5], students with active learning styles expected to show behavioral patterns of online learning as follows: (i) frequently send messages in the forum to ask questions and explain things, (ii) perform more exercise due to preference of testing and trying things out and, (iii) spend only little time on reading material like content object because of favour to retain information by doing something active with learned material. On the other hand, students with reflective learning style expected to show behavioral patterns of online learning as follows: (i) participate passively by frequently reading the postings and rarely posting messages in the forum due to the preference of thinking about the material and work alone, (ii) visit and spend more time on reading material to think and reflect about the material and (iii) Take longer time on exercise for reflecting on their result but expected to perform less exercise. After determining the relevant patterns of behaviour based on Felder and Silverman learning style model dimensions, threshold value is needed to classify the occurrence of behaviour pattern.

3.2. Threshold Value

The previous section describes the relevant patterns of behavior for information processing dimension while using LMS. Hence, in order to classify the occurrence behaviour patterns, a 4-item scale is used to divide the behaviour into four groups: high, moderate, low and no information occurrence. Classification based on the threshold values. The threshold value is used to determine whether the presence of behavior pattern indicators is high (3), moderate (2), low (1) or no information about students’ behavior (0) with respective learning style.
There are two methods that can be used to determine the threshold value which are the average of behavior patterns and the general threshold value. However, the use of the average value method is not suitable for small or medium group as it is based on the pre-determined distribution [32]. On the other hand, the use of the general threshold value method is based on the values found in the previous studies and need to be adjusted according to the structure of the course, the subjects of the course and the experience of the students [34, 35]. Table 1 shows the behavior patterns those are relevant to the features of Felder and Silverman learning styles model along with the proposed threshold value [26, 36].

Table 1. Behaviour pattern based on learning style dimension of Felder and Silverman

<table>
<thead>
<tr>
<th>Features</th>
<th>Behaviour Pattern</th>
<th>Pattern Description</th>
<th>Threshold value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
<td>content_visit</td>
<td>percentage of visited content object (based on the number of available content)</td>
<td>75% 100%</td>
</tr>
<tr>
<td></td>
<td>content_stay</td>
<td>Percentage of time spent on the content object (based on a predefined expected value)</td>
<td>50% 75%</td>
</tr>
<tr>
<td>Exercise</td>
<td>exercise_visit</td>
<td>percentage of performed exercises (based on the number of available exercises)</td>
<td>25% 75%</td>
</tr>
<tr>
<td></td>
<td>exercise_stay</td>
<td>Percentage of time spent on exercises (based on a predefined expected value)</td>
<td>50% 75%</td>
</tr>
<tr>
<td>Forum</td>
<td>forum_visit</td>
<td>number of visit forum (per week)</td>
<td>7 50</td>
</tr>
<tr>
<td></td>
<td>forum_stay</td>
<td>time spent in the forum (per week)</td>
<td>5 min 30 min</td>
</tr>
<tr>
<td></td>
<td>forum_post</td>
<td>number of postings in forum (per week)</td>
<td>1 10</td>
</tr>
</tbody>
</table>

4. Methodology

To obtain the threshold value, the behavior patterns of 20 electrical engineering students who were taking the interactive multimedia subject in their semester five were analyzed. The analysis took over 11 weeks of the whole learning process of a semester that covered up the 12 major sub-topics on the subject. The teaching of the subject utilized three types of learning materials: (1) learning notes in the SCORM format (12), exercises (6) and forums (12).

First, learning notes in the SCORM format was used to obtain the threshold value of content_stay and content_visit behavioral patterns. Such format was chosen because of its ability to track the behavior patterns of each linked concept and it could provide a clearer and a systematic time period. The learning notes are made up of text and image elements where the students can only read the notes online and notes printing are not allowed. Next, the exercise activities were used to obtain the threshold value of exercise_stay and exercise_visit behavioral patterns. For each of the two sub-topics, students were given the freedom to do the exercise to test their understanding of the learning notes provided. Finally, to obtain a threshold value of the forum_visit, forum_stay and forum_post behavior patterns, the activities in the online forum which included the problem-solving discussions, expressing opinions and open discussions on specific social problems were analyzed.

The raw data of each student’s behavior pattern were obtained from their log data while performing the online learning. Next, the behavior patterns of all students were analyzed statistically to obtain the threshold value based on the percentile value of the cut point of three equal groups. The second percentile value was used to determine the first threshold value and the third percentile value was used to determine the second threshold value. And lastly, the threshold value obtained will be compared with the previous study. The justification for selecting the threshold
value was based on the consideration of the actual behavior patterns of students for each characteristic studied in the course.

5. Results and Discussion

According to Alberer et al., [34] and Roblyer and Wiencke [35], the threshold value varies to different courses because it depends on the course structure, subject and students’ experience of using the learning materials and activities. Accordingly, based on the threshold value suggested by earlier researchers such as Rovai and Barnum [36] and Garcia et al. [26], therefore, this study coordinated the threshold value based on the characteristics of the course by taking into consideration the actual use of provided learning materials and activities. Based on Table 2, the threshold value for content_visit is 60% and 80%. When compared the value to the threshold value proposed by Garcia et al. [26] and Atman et al. [33] in Table 3, it was found that the second threshold value which was 80% did not reach 100% or more than 100% because in this study, the content object was taught during face-to-face learning sessions. This situation caused many students to visit content object in order to find information while doing exercises and discussing in the forum. As a result, the use of object content had resulted in low threshold value for content_stay compared to the study done by Garcia et al. [26], Graf et al. [5] and Atman et al. [33], which are 20% and 45%.

Table 2. Thresholds for determining behaviour pattern

<table>
<thead>
<tr>
<th>Behaviour Pattern</th>
<th>Thresholds</th>
</tr>
</thead>
<tbody>
<tr>
<td>content_visit</td>
<td>60% 80%</td>
</tr>
<tr>
<td>content_stay</td>
<td>20% 45%</td>
</tr>
<tr>
<td>exercise_visit</td>
<td>60% 80%</td>
</tr>
<tr>
<td>exercise_stay</td>
<td>30% 50%</td>
</tr>
<tr>
<td>forum_visit</td>
<td>8 12</td>
</tr>
<tr>
<td>forum_stay</td>
<td>5min 8min</td>
</tr>
<tr>
<td>forum_post</td>
<td>1 2</td>
</tr>
</tbody>
</table>

Table 3. Threshold value based on literature

<table>
<thead>
<tr>
<th>Behaviour Pattern</th>
<th>Rovai and Barnum</th>
<th>Garcia et al.</th>
<th>Graf et al.</th>
<th>Atman et al.</th>
</tr>
</thead>
<tbody>
<tr>
<td>content_visit</td>
<td>-</td>
<td>75% 100%</td>
<td>10% 20%</td>
<td>75% 125%</td>
</tr>
<tr>
<td>content_stay</td>
<td>-</td>
<td>50% 75%</td>
<td>50% 75%</td>
<td>75% 100%</td>
</tr>
<tr>
<td>exercise_visit</td>
<td>-</td>
<td>25% 75%</td>
<td>25% 75%</td>
<td>50% 75%</td>
</tr>
<tr>
<td>exercise_stay</td>
<td>-</td>
<td>50% 75%</td>
<td>50% 75%</td>
<td>25% 50%</td>
</tr>
<tr>
<td>forum_visit</td>
<td>7 50</td>
<td>7 14</td>
<td>75% 100%</td>
<td></td>
</tr>
<tr>
<td>forum_stay</td>
<td>-</td>
<td>1min 2min</td>
<td>50% 75%</td>
<td></td>
</tr>
<tr>
<td>forum_post</td>
<td>1 10</td>
<td>-</td>
<td>1 2</td>
<td></td>
</tr>
</tbody>
</table>

Next, referring to the threshold value for behavior patterns of visiting exercise, this study used a quite similar value to Atman et al. [33]. As the threshold values were obtained based on the real behavior patterns of students, it was found that many students did around 60% to 80% of the total exercise provided. Garcia et al. [26] and Graf et al. [5] suggested the values of 25% and 75% of the threshold value of exercise_visit behavior patterns. Yet, for the exercise_stay threshold value, the value of 30% and 50% had been chosen as referring to the given type of question. The type of questions given is the multiple-choice type with short and easily understood questions. However, Based on the actual behavior patterns, it is found that most students need shorter time from the expected time.

The last is the threshold value for the behavioral patterns in the forum discussions. Based on the threshold value proposed by Rovai and Barnum [36], 50 or more visits per week and 10 forum posts are considered high and there is no value for the time spent on the forums. However, according to Graf et al. [5] there are studies that claimed the
proposed threshold value is in the range of 30 minutes per week. On the actual study of the actual behavior, Graf et al. [8] suggested a value of 5 and 10 minutes per week. Based on this study, the threshold value for forum_stay was almost similar which was at 5 to 8 minutes per week. In regards to the observations, most of the students allocated their time just to read the received messages rather than making new posts. For this reason, the threshold value for forum_post are 1 and 2. While the threshold value for the visit the forum behavior patterns are 8 and 12 per week.

The findings of this research support on the need to determine the threshold value of every pattern for identifying students’ learning style. Even though there are recommended threshold value based on literature, these threshold values can change with respect to the consideration of actual usage of the investigated pattern in the course.

6. Conclusion

Identifying students’ learning styles automatically through their actual behavior patterns while using the LMS plays an important role in pedagogy aspect. The knowledge about students’ learning styles can be used to help students learn effectively by matching the learning materials and activities with their learning style or known as adaptive learning. The use of literature-based approach which is practical and easy is believed to help most instructors who use the LMS to identify their students' learning styles. However, the threshold value should be determined based on the actual use of LMS to ensure that these values are consistent with the structure of the courses used in the LMS and suit the student experience.

References


