Toward Integrating the Pedagogical Dimension in Automatic Learner Modeling within E-Learning Systems

Mohamed Koutheair KHRIBI 1, Mohamed JEMNI 1, Olfa NASRAOUI 2
1 Technologies of Information and Communication Lab,
Higher School of Sciences and Technologies of Tunis, University of Tunis
5, Av. Taha Hussein, B.P. 56, Bab Mnara 1080, Tunis, TUNISIA
mk.khribi@uvt.rnu.tn, mohamed.jemni@fst.rn.tn
2 Knowledge Discovery & Web Mining Lab,
Dept. of Computer Engineering & Computer Science,
Speed School of Engineering,
University of Louisville, Louisville KY 40292, USA
olfa.nasraoui@louisville.edu

Abstract

In order to automatically provide the most appropriate learning objects to e-learners, a special interest should be given to the process of building the learners’ models. First, we need to identify which relevant information to include in the learner’s model, while taking into account various pedagogical considerations, and second how to accurately infer the learners’ preferences and characteristics from their online behavior and activities. This paper presents a preliminary study of the possibility of integrating educational preferences in the learner’s model and detecting them automatically within e-learning systems.

1. Introduction

Despite the availability of abundant educational resources and services, it is still difficult to decide which learning objects better match the student’s needs in a given situation, unless we accurately know the profile of that particular learner. The learner’s profile or model is generally built based on a range of information, gathered through the user’s implicit and/or explicit feedback, describing the learner’s preferences. Features that are related to pedagogical aspects have to be taken into account when modeling learners, such as educational preferences, learning styles, skills, etc. [2]. Basically, two main student modeling approaches can be outlined: collaborative student modeling and automatic student modeling. The collaborative student modeling approach requires students to provide explicit information about their preferences and needs. In the automatic student modeling approach, gathering information is done rather automatically based on the online behavior and activities of students. Automated building of learner models involves the automated detection of all basic information composing the model including the educational preferences and learning styles. In this paper, we are concerned with this issue which gives a complementary and fruitful pedagogical dimension to the automate recommender system that we have already designed to personalize e-learning based on Web clickstreams [9]. In the following section, we identify essential components to include in the proposed learner’s model. In section 3, we focus on the process of automated student modeling. Section 4 introduces the phase of group modeling. Finally, section 5 presents our conclusions and future work.

2. Components of a Learner’s Model

Common types of information used in learner models include the learner’s knowledge, demographic information, preferences, etc. These components are strongly connected to the application of the learner model. According to Brusilovsky [3], there are two main categories: domain specific information and domain independent information. In the following, we adopt this categorization but further split the learner model into three main components as follows:
- The learner profile containing general student information such as Identification data, and demographic information;
- The learner’s knowledge model storing knowledge and behaviour of the student exploring learning objects and services within the e-learning system. It can be represented by a sequence of weighted visited learning
objects i.e. a vector of visited learning objects or curriculum elements in which the student was interested.

- The learner’s educational preferences containing educational attributes and the learning style. A range of elements belonging to the LOM (Learning Objects Metadata) category description can be considered such as “the learning resource type” to describe the learner’s preferences on material. This attribute for example can take as values: {Exercise, Simulation, Questionnaire, Exam, Experiment, Self assessment, Lecture, Slide}. Concerning the learning style component, we adopt the Felder-Silverman learning styles model (FSLSM) in which, learning styles models are described in detail by characterizing each learner according to four dimensions: active/reflective, sensing/intuitive, visual/verbal, sequential/global [5].

3. Automated Learner Modeling

3.1 Learner Knowledge Model Construction

We apply Web usage mining techniques to analyze user sessions. The input data for this step consists of clickstream data stored in Web server access log files and/or visit statistics stored dynamically in a database. The output of these steps, for each student, is a set of sessions gathered over a period of time. Let \( L \) be a set of visited learning objects: \( L = \{ L_1, L_2, ..., L_m \} \), and let \( L \) be a set of \( m \) learners registered in a specific course within an e-learning environment, \( L = \{ L_1, L_2, ..., L_m \} \); the learner knowledge model \( L_K \), corresponding to the learner \( L_i \in L \) is represented by a set of \( p \) sessions: \( S^p_i \), where each \( S^p_i \) is a subset of \( k \) weighted visited learning objects:

\[
S^p_i = \langle (L_{O_1}^i, w(L_{O_1}^i)), (L_{O_2}^i, w(L_{O_2}^i)), (L_{O_3}^i, w(L_{O_3}^i)), ..., (L_{O_k}^i, w(L_{O_k}^i)) \rangle,
\]

where each \( L_{O_l}^i \) is a subset of \( L \) for some \( l \in \{1, ..., n\} \), and \( w(L_{O_l}^i) \) is the weight associated with learning object reference \( L_{O_l}^i \) in the session \( S^p_i \) corresponding to the \( i \)th student \( L_i \). The knowledge model of a learner \( L_i \) can be represented by a matrix \( M(p, n) \) where \( p \) is the number of completed sessions and \( n \) the cardinality of unique visited learning objects:

\[
\begin{pmatrix}
    w(L_{O_1}^p) & \cdots & 0 \\
    \vdots & \ddots & \vdots \\
    0 & \cdots & w(L_{O_n}^p)
\end{pmatrix}
\]

3.2 Educational preference detection

3.2.1 Detection of educational attributes

First, we detect educational attributes for all visited learning objects \( L \). Educational attributes can be located in the XML files describing the \( L \) (we assume here that only Standard defined Learning Objects are considered). These attributes are detected using automated crawling, parsing, and indexing techniques (as done by search engines). The occurrence of the educational attributes values over all visited learning objects can show us the \( L \) that the learner may prefer.

3.2.2 Toward learning style prediction

Recent research has dealt with the automated detection of learning styles. In [6], Bayesian networks were used to detect and model student learning styles. The approach was evaluated by comparing the proposed Bayesian model with the results obtained using the ILS questionnaire. [4] investigated the use of Decision Trees and Hidden Markov Models, while [7] studied the behavior of students during their online work to gather hints about their leaning styles. Using a simple rule based mechanism, learning styles were calculated using the gathered indications. The evaluation of the approach demonstrated good results and suitability for identifying learning styles with respect to FSLSM. Generally, the automated detection of learning styles is accomplished in two phases (1) determining relevant learners’ behavior for the detection process, and (2) inferring learning styles from behavior. The second part can be resolved through two different approaches, a data-driven approach and a literature-based approach [8]. Defining patterns, to determine relevant behaviors, is done based on two assumptions, first, derived patterns from literature must be relevant in the identification of learning styles, second, tracking considered patterns must be feasible through e-learning systems. Once patterns are well defined, learning styles should be extracted from tracked data. We intend to apply patterns proposed in the literature, and easy to track in the known e-learning systems. Eight classes can be used since there are eight learning styles (active, reflective, sensing, intuitive, visual, verbal, sequential and global). Each class can be defined by a set of \textit{if-then} rules. Then, based on each student tracked data corresponding to the \( LS \) pattern, the student’s learning style will be predicted.

---

1 http://www.engr.ncsu.edu/learningstyles/ilsweb.html
4. Group Modeling

Once the learners’ models have been estimated, we apply a three-level model based collaborative filtering approach. In the first and second level, we apply clustering techniques based on similarities and dissimilarities among student educational preferences. In the third level, association rules are discovered from each cluster. The first level of student session clustering is achieved based only on their educational preferences (i.e., considered educational attributes and learning style) as defined previously. We obtain then different clusters of sessions, each representing learners having similar educational preferences. In the second level, clustering is applied again on each of the previously obtained clusters, but this time using similarities or dissimilarities between the preferred visited learning objects. A variety of clustering techniques can be used for clustering sessions. Regardless of which method is used, the first clustering level will result in a set \( C = \{C_1, C_2, \ldots, C_{|C|}\} \) of clusters. The second clustering level applied on each obtained cluster \( C_i \) will result in a set \( C_i = \{C_i^1, C_i^2, \ldots, C_i^{|C_i|}\} \), where each cluster \( C_i^j \) is a subset of student sessions representing a group of similar learners with similar access patterns.

Each obtained cluster \( C_i^j \) is split into preferred learning object sets using a frequent itemset mining algorithm that extracts frequently co-occurring LO sets in sessions belonging to each cluster. We use the discovered cluster association rules to accomplish this task. These association rules (AR) capture the relationships among the LO based on their co-occurrence across sessions. AR discovery methods such as the Apriori algorithm [1] can be used directly.

5. Conclusions

In this paper, we described an approach to model students automatically while taking into account their educational preferences. We have applied web mining techniques to obtain accurate sessions beyond the student navigation history and to automatically extract the educational attributes of the visited learning objects. The learner’s model is composed of three components: learner’s profile, learner’s knowledge and learner’s educational preferences. Two main elements describe the last component: educational attributes and learning style. After construction of the student models, we build group models using a three-level collaborative modeling approach. In the future, we plan to evaluate the proposed modeling approach, and to integrate the student modeling module within a recommender system for e-learning environments.

6. References