RECOMMENDING LEARNING MATERIAL IN INTELLIGENT TUTORING SYSTEMS

Jarosław Bernacki

1 Department of Computer Science
Wroclaw University of Technology

KEY WORDS: Centre of Information Technologies, information, technologies

ABSTRACT: Nowadays, intelligent e-learning systems which can adapt to learner's needs and preferences, became very popular. Many studies have demonstrated that such systems can increase the effects of learning. However, providing adaptability requires consideration of many factors. The main problems concern user modeling and personalization, collaborative learning, determining and modifying learning scenarios, analyzing learner's learning styles. Determining the optimal learning scenario adapted to students' needs is very important part of an e-learning system. According to psychological research, learning path should follow the students' needs, such as (i.a.): content, level of difficulty or presentation version. Optimal learning path can allow for easier and faster gaining of knowledge. In this paper an overview of methods for recommending learning material is presented. Moreover, a method for determining a learning scenario in Intelligent Tutoring Systems is proposed. For this purpose, an Analytic Hierarchy Process (AHP) method is used.

1. INTRODUCTION

Intelligent Tutoring Systems (ITS) are computer software that can conduct learning process without physical teacher [19]. The most important goal of Intelligent Tutoring Systems is to provide an optimal learning path to learners. Many research showed that ITS can successfully increase learning effects, for example, by creating of collaborative learning groups [1].

In this paper a method for determining a personalized learning scenario is proposed. According to pedagogical/psychological research, some factors that have a strong influence on attractiveness of learning scenario, are: the content of scenario lessons, the level of difficulty of lessons and the preferred version of lessons. Preferred version of lessons can be determined by many learning styles questionnaires, i.e. Paragon [21], Memletics [17] or Felder / Silverman [8]. In this work, a personalized learning scenario is proposed with use of Analytic Hierarchy Process (AHP) method. In this approach, a student is a user of intelligent e-learning system, where some learning scenarios are presented. Each scenario is a graph-based set of lessons, at particular level of difficulty and preferred form of presentation, based on Felder / Silverman model (text, graphic or
interactive). For each pair of scenarios, student gives the score of preferring a given scenario (based on a scale proposed by Saaty [22]). Then, a ranking of scenarios is calculated by AHP method. The scenario with the highest rank value is recommended to a student.

The paper is organized as follows: section 2 contains an overview of methods for recommending a learning material, in section 3 there are presented some basic notions. Section 4 presents a method for determining a learning scenario and section 5 concludes this work with some possible future directions.

2. RECOMMENDING LEARNING MATERIAL

The problem of providing personalized learning scenario is widely discussed in literature. The majority of works prove that the most important personalization parameter is learner's level of knowledge [20], [6]. Other parameters can be: learning goals [25], [18], media preferences [16], [3] or language preference [6]. A comprehensive overview of personalized systems is presented in [7]. Overview of essential content of user profile for providing adaptability is presented in [5].

In [7] the personalization was made by asking students in questionnaire about their preferences. Then, personalized scenarios were recommended based on satisfaction rates, given by each student. However, this approach has some disadvantages. For example, if student does not vote in accordance with truth, it may generate wrong recommendations. Moreover, voting student can be not objective.

In [11], a method for recommendation URL links, based on mining techniques and content-based profiles is proposed. Each learner should be registered in intelligent learning system and has a profile. This profile is build using student's activity (information is stored in system log les). Then, based on the analysis of these log les, the recommendations are generated in content-based way. For this purpose, Association Rules (ARs) techniques (apriori algorithm) are used.

Paper [9] presents a method which combines peer-learning and social learning theories to produce recommendations. The used idea is based on recommending learning materials with similar content and indicating the quality of learning materials based on good learners' ratings. The objective is to recommend the learning materials that are similar to those of the viewing item. Then, learners are selecting learning resources. Each material is rated by good learners. Good learners are learners who have studied the learning resources and completed knowledge test with the score not less than 80%.

Kind of hybrid method is proposed in [23]. It combines sequential pattern mining and multidimensional attribute-based collaborative filtering (CF). The sequential pattern based approach uses the apriori and PrefixSpan algorithms to discover patterns in accessing material and to generate recommendations. Learner Preference Tree (LPT) is used to take into account multidimensional attribute of materials and learners' ratings. LPT model considers access time-length of a material and visiting frequency of a material, and then estimates its importance for a learner.
Content-based recommendation using clustering methods is described in [2]. Recommendations contain also some social connections that could improve the learning process. A clustering technique is used in order to recommend both didactic resources and learning groups and to facilitate the learning process. Items (contents, users, groups) are grouped in clusters using the Shared Near Neighbor algorithm, and the recommendation method examines the proximity of each item to the clusters. Then, it builds the list of recommended items.

In work [24] an ontology-based recommendation method is proposed. As the main point for recommendation, the personalization is considered. Authors propose to recommend learning materials according to learner's learning styles. Authors propose to use the well-known Felder / Silverman's model and also Kolb's model. The ontology classify learners according to their learning styles and recommend appropriate learning resources. The SPARQL query language is used and the simulation of recommendation process is implemented in Protegé Software.

In [26] a comprehensive analysis of recommending learning resources according to Felder / Silverman's learning styles is presented. Authors in detail examine, which kind of learning material is appropriate for a given learning style. Also a method for recommendation is proposed.

In [12] a system called Protus is described. Its main aim is to generate recommendations based on personalization parameters, which are: learning styles (and learners' habits), learners' interests and their level of knowledge. The recommendation process consists of two phases: recognizing learning styles, and analyzing learners' interests and habits. The phase of analyzing interests and habits is performed with an AprioriAll algorithm, which is used to generate the final recommendations.

In [10] there is proposed a method for personalized adaptive English learning under fuzzy logic, AHP method and Item Response Theory (IRT). Learner's skills are modeled by a fuzzy trapezoidal membership function. Difficulty of learning material is estimated by one-parameter IRT model. AHP method is used to calculate the recommendation score, based on the suitability for the learner.

According to [15], learning process should take into account student's individual needs and learning styles. Moreover, intelligent learning environment should be able to adapt and modify own behaviour to follow the expectations of the student. In this paper, personalized learning path is provided with use of the fuzzy matching rules. Both student's profile and learning content are represented as vectors. A fuzzy matching rule assigns suitable learning material to a student.

Personalization based on Item Response Theory is presented in [4]. Considered is both difficulty of learning material and learner abilities. Course materials are estimated by item characteristic function proposed by Rasch with single difficulty parameter. Learner abilities are modeled by maximum likelihood estimation (MLE). Recommendation is prepared by recommendation agent that uses the information's about material's difficulty and learner's abilities.
3. A METHOD FOR LEARNING CONTENT RECOMMENDATION

In this section we introduce a method for a learning material (scenario) recommendation. This method is based on a graph-based structure of knowledge which represents learning scenarios.

3.1. Structure Of Knowledge And Definitions

The used structure of knowledge is a modification of structure presented in works [13] [14]. Let\( P \) denote a finite set of lesson.

**Definition 1.** A lesson \( p_i \in P \), where \( i \in \{0, \ldots, q\} \) is a set of lesson versions \( v_{r,k}^{(i)} \in p_i \) at a certain level of difficulty \( r \), where \( r \in \{1, \ldots, d\} \) and \( k \in \{1, \ldots, m\} \); \( d \) - the cardinality of set of levels of difficulty. Each lesson occurs in \( m \) versions; \( q \) denotes the cardinality of set of lessons.

**Definition 2.** Let \( R_c \) denote a set of linear orders on a set \( P \). Relation \( \alpha \) is linear, if it is reflexive, transitive, antisymmetric and consistent.

**Definition 3.** A graph structure of knowledge is a directed and labeled graph:

\[ G = (P, E, \mu) \]

where: \( P \) - set of vertices (lessons); \( E \) - set of edges, \( E \subseteq P \times P \); \( \mu : E \rightarrow L \) - a function assigning labels to the edges, \( L \) - set of labels: \( L = \bigcup_{f=1}^{\text{card}(R_c)} a_f \), where, \( a \in R_c \).

**Definition 4.** A Hamiltonian path \( hp \) based on order \( a \in R_c \) in graph \( Gr \) is a sequence of vertices (lessons):

\[ hp = < p_0, \ldots, p_q > \]

where for all \( i \in \{0, \ldots, q\} \), \( p_i \neq p_{i+1} \)

**Definition 5.** A scenario \( s \) is a Hamiltonian path \( hp \) (based on order \( a \in R_c \)), in which there is exactly one element of each set \( p_i \), \( i \in \{0, \ldots, q\} \):

\[ s = < d_{r,k}^{(0)}, \ldots, d_{u,n}^{(q)} > \]

where \( r, u \in \{1, \ldots, d\} \) and \( k, n \in \{1, \ldots, m\} \).

**Example 1.** Some example scenarios generated on structure presented in a Figure 1:

\[ s_1 = < d_{2,3}^{(1)}, d_{2,1}^{(2)}, d_{1,3}^{(3)}, d_{1,1}^{(4)} > \]

\[ s_2 = < d_{1,2}^{(1)}, d_{2,3}^{(4)}, d_{2,2}^{(2)}, d_{1,3}^{(3)} > \]
**Definition 6.** An edit distance is a measure for determining a difference between two sequences. Levenshtein distance is a number of operations (insertion, deletion and exchange) that are essential to transform one sequence to another. If we consider two scenarios as a sequence of characters, the distance between them could be calculated with use of a Levenshtein distance and then normalized to, for example interval $[0, 1]$.

**Example 2.** Suppose we have two following scenarios:

$s_1 = d_{3,3}^{(0)}, d_{1,2}^{(1)}, d_{5,2}^{(2)}, d_{2,3}^{(3)}, d_{1,3}^{(4)}, d_{4,1}^{(5)}, d_{2,2}^{(6)}$

$s_2 = d_{5,2}^{(0)}, d_{4,1}^{(6)}, d_{1,3}^{(5)}, d_{2,1}^{(1)}, d_{1,3}^{(2)}, d_{5,2}^{(3)}, d_{4,1}^{(4)}$

Consider them as a sequence of characters. The Levenshtein number of operations to transform $s_1$ into $s_2$ equals 6.

Next, we can use min-max normalization to represent this value in an interval $[0, 1]$ with $new_{\min} = 0$ and $new_{\max} = 1$, which is defined as follows:

$$V' = \frac{(V - \text{min})}{\text{max} - \text{min}} (new_{\max} - new_{\min}) + new_{\min}$$

where:
- $V$ - analyzed value;
- $\text{min}$ - minimal value in input;
- $\text{max}$ - maximal value in input;
- $new_{\min}$ - new minimal
- $new_{\max}$ - new maximal value.

We obtain:

$$V' = \frac{6 - 0}{7 - 0} \cdot (1 - 0) + 0 = \frac{6}{7} = 0.86$$

Therefore, the measure of dissimilarity between scenarios $s_1$ and $s_2$ equals 0.86.
3.2. Determining A Learning Scenario

A problem of determining an optimal learning scenario can be solved by using an Analytic Hierarchy Process (AHP) method. It is an algorithm widely used in group decision making that allows selecting decisions that meet the preferences of decision-maker. In this method, a model of a given problem is built as a hierarchy, given by expert or decision-maker. Then, an expert (or decision-maker) gives the scores for pair-wise comparisons for elements of the hierarchy, which generates so-called matrix of preferences (or matrix evaluations). Based on such matrix, there are calculated so-called priorities representing the validity of the element from hierarchy. Decision is a choice of options sorted as a ranking.

Table 1 The AHP scale for pair-wise comparisons

<table>
<thead>
<tr>
<th>Value</th>
<th>Rating of A to B</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>A is extremely preffered</td>
</tr>
<tr>
<td>7</td>
<td>A is very strongly preffered</td>
</tr>
<tr>
<td>5</td>
<td>A is strongly preffered</td>
</tr>
<tr>
<td>3</td>
<td>A is weakly preffered</td>
</tr>
<tr>
<td>1</td>
<td>A is equivalent to B</td>
</tr>
</tbody>
</table>

In this approach, AHP method could be used for choosing a scenario. This can be realized in a following way:

- A number of scenarios are presented to a student;
- Student compares pairs of scenarios and gives the scores for each scenario as a numeric value presented in Table 1;
- A preference vector (based on given scores) is calculated by AHP method;
- The maximum value in the preference vector responds to the optimal scenario.

The comparison of given elements are given by numeric values, where the scale of values is from 1 to 9. This scale is presented in Table 1.

Formally, determining a learning scenario with an AHP method can be described as follows: suppose we have n learning scenarios $s_1, s_2, \ldots, s_n$. We want to find a scenario $s^*$ that has the greatest validity of scenarios hierarchy.

First, we need to generate a matrix of preferences of given scenarios:
For example, the relation between scenario $s_1$ and $s_2$ is $u_1$, and so on.

Table 2. Random Consistency Index $RI$

<table>
<thead>
<tr>
<th>$n$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$RI$</td>
<td>0</td>
<td>0.58</td>
<td>0.9</td>
<td>1.12</td>
<td>1.24</td>
<td>1.32</td>
<td>1.41</td>
<td>1.45</td>
<td>1.51</td>
</tr>
</tbody>
</table>

The next step is to calculate sums of elements in columns in preference matrix $M^{(0)}$:

$$c_i^{(0)} = \sum_{l=1}^{k} m_{il}^{(0)}$$

(4)

Next, a normalized matrix $M_i^{(0)}$ is calculated as a quotients of elements of preference matrix divided by sum of elements in given column:

$$\overline{m}_{il}^{(0)} = \frac{m_{il}^{(0)}}{\sum_{l=1}^{k} m_{il}^{(0)}}$$

(5)

Then, average values of normalized matrix $\overline{M}^{(0)}$ are calculated as:

$$\overline{s}_l^{(0)} = \frac{1}{k} \sum_{i=1}^{k} m_{il}^{(0)}$$

(6)

These average values determine the preference vector:

$$\overline{s}^{(0)} = [\overline{s}_1^{(0)}, \overline{s}_2^{(0)}, \ldots, \overline{s}_l^{(0)}]$$

The last step is to check consistency of preference matrix $M^{(0)}$

$$\lambda_{max}^{(0)} = \sum_{i=1}^{n} c_i^{(0)} \cdot \overline{s}_i^{(0)}$$

The $\lambda_{max}^{(0)}$ value is used for determining a Consistency Index ($CI$) ratio: $CI = \frac{\lambda_{max}^{(0)} - n}{n-1}$. Then, the Consistency Ratio ($CR$) is calculated as a comparison of Consistency Index and Random Consistency Index ($RI$): $CR = \frac{CI}{RI}$.

The $RI$ values are presented in Table 2 ($n$ is a number of compared items in preference matrix).

If the $CR$ value is smaller than 0.1 (10%) then the preference matrix $M^{(0)}$ is consistent. Otherwise, the preferences should be redefined. There are some methods in literature for delivering the consistency of preference matrix, for example: the greatest eigenvalue, least squares method or method of least logarithmic squares.
This approach can be described as a following pseudocode:

Algorithm 1 (Scenario recommendation).

Input: Scenarios: $s_1, s_2, \ldots, s_n$

Output: Vector of preference $\mathbf{s}^{(0)}$

foreach pair of scenarios assign score do
| generate matrix of preferences (scenarios comparison) $\mathbf{M}^{(0)}$

End

foreach column in matrix $\mathbf{M}^{(0)}$ do
| calculate sum of elements in columns: $c^{(0)}_l = \sum_{k=1}^n m^{(0)}_{lk}$

End

foreach element in matrix $\mathbf{M}^{(0)}$ do
| calculate normalized matrix $\mathbf{\bar{M}}^{(0)}_{il} = \frac{m^{(0)}_{il}}{\sum_{l=1}^n m^{(0)}_{il}}$

End

foreach row in normalized matrix $\mathbf{\bar{M}}^{(0)}$ do
| calculate average of elements in rows: $s^{(0)}_l = \frac{1}{\sum_{k=1}^n m^{(0)}_{lk}}$

End

Check consistency of matrix $\mathbf{\bar{M}}^{(0)}$:
\[ \lambda_{max}^{(0)} = \sum_{l=1}^n c^{(0)}_l \bar{s}^{(0)}_l \]

Determine preference vector: $\mathbf{s}^{(0)} = \left[ s^{(0)}_1, s^{(0)}_2, \ldots, s^{(0)}_l \right]$

Determine arg max $\mathbf{s}^{(0)}$

The computational complexity of the Algorithm 1 is $O(n^2)$.

Example 1. Consider that there are three scenarios $s_1, s_2, s_3$. The content of the scenarios is presented to a student, for instance on website of some e-learning system. Each scenario consist of set of lessons, on particular level of difficulty and version of lesson (text, graphic or interactive). Suppose that student gave the following scores in pair comparisons:

- $(s_1, s_2) \rightarrow 5$ - the $s_1$ scenario is strongly preferred than scenario $s_2$;
- $(s_1, s_3) \rightarrow 7$ - the $s_1$ scenario is very strongly preferred than scenario $s_3$;
- $(s_2, s_3) \rightarrow 3$ - the $s_2$ scenario is weakly preferred than scenario $s_3$.

Then, the matrix of scenarios comparisons is defined on above scores as follows:

\[
\mathbf{M}^{(0)} = \begin{bmatrix}
1 & 5 & 7 \\
1 & 5 & 1 \\
1 & 1 & 3 \\
7 & 3 & 1
\end{bmatrix}
\]
Next, we calculate sums of columns in matrix $M^{(0)}$:

\begin{align*}
    c_1^{(0)} &= 1.34 \\
    c_2^{(0)} &= 6.33 \\
    c_3^{(0)} &= 11
\end{align*}

In the next step, we calculate a normalized matrix $M^{(0)}$ and we obtain:

$$
    \overline{M}^{(0)} = \begin{bmatrix}
        0.75 & 0.79 & 0.64 \\
        0.15 & 0.16 & 0.27 \\
        0.11 & 0.05 & 0.09
    \end{bmatrix}
$$

Another step is to calculate the average values of rows in matrix $\overline{M}^{(0)}$ and to determine a preference vector:

\begin{align*}
    \overline{s}_1^{(0)} &= 0.73 \\
    \overline{s}_2^{(0)} &= 0.19 \\
    \overline{s}_3^{(0)} &= 0.08
\end{align*}

\[ \overline{s}^{(0)} = [0.73; 0.19; 0.08] \]

We choose the arg max $\overline{s}^{(0)} = 0.73$ that stands for the scenario $s_1$.

The last step is to check the consistency of matrix of comparisons $M^{(0)}$: $\lambda_{max}^{(0)} = 3.06$

The CI index equals: $\frac{3.06 - 3}{3-1} = 0.03$ and the CR index $= \frac{0.03}{0.58} = 0.05 < 0.1$. This denotes that the matrix of scenarios comparisons $M^{(0)}$ is consistent. Therefore, a scenario $s_1$ is optimal and is recommended to a student.

One of advantages of this method is fully personalization, because student decides on her/his own, what interest her/him most. Moreover, typical recommender systems, based on content-based (CB) or collaborative filtering (CF), grapple with a cold start problem. This situation happens, when in system there is a small number of users – it is quite difficult to propose an appropriate recommendation and even sometimes recommendations are generated in a random way. Another advantage is that AHP-based system works in real time.

4. CONCLUSION

In this paper a method for determining a learning scenario is proposed. For generating a learning scenario, the following factors are considered: the content of a scenario lessons, level of its difficulty and the presentation version, based on Felder / Silverman learning styles model. An Analytic Hierarchy Process algorithm is used to determine the best scenario. Proposed method has many advantages. One of them is personalization - student makes her/his own recommendation for learning scenario and there is no need to collect information’s about other students to propose a recommendation, as it is done in for example collaborative filtering (CF) systems.
In future works it is planned implement a prototype of an e-learning system and to conduct an experiment that could check the efficiency of this method on a representative group of students. Moreover, a method for delivering the consistency of preference matrix should be considered.

Another problem is to try to use this method for collaborative learning groups, where group profile should be determined.

LITERATURE

[21] Paragon Learning Style Inventory: http://web.calstatela.edu/faculty/jshindl/plsi/ (last access March 10, 2015)