Toward a fully automatic learner model based on web usage mining with respect to educational preferences and learning styles

Mohamed Koutheaïr KHRIBI
University of Tunis
mk.khribi@uvt.rnu.tn

Mohamed JEMNI
University of Tunis
mohamed.jemni@fst.rnu.tn

Olfa NASRAOUI
University of Louisville
olfa.nasraoui@louisville.edu

Sabine GRAF
Athabasca University
sabineg@athabascau.ca

Kinshuk
Athabasca University
kinshuk@athabascau.ca
Learner Modelling Approaches

- Collaborative
- Automatic

- Advantages of automatic learner modeling :
  - No additional work for learners ;
  - Uses information from a time span; higher tolerance
  - Allows dynamic updating of information
General Aim

Building an automatic recommendation system for Learning Management Systems.
Research Question

How to automatically model learners and groups of learners based on implicit data from their interactions and online activities in the learning management system, taking into account the learners’

* educational preferences
* learning styles
General Features of the Proposed Approach

- Automatic, dynamic and based solely on learner usage sessions;
- Input data is composed from collected implicit data tracked and saved in LMS database and/or web server log files;
- Based on web usage mining techniques;
- Educational preferences and learning styles are considered and identified automatically.
## Proposed Learner Model

The proposed learner model components are defined as:

\[ LM_i = \{PR_i, LK_i, LEP_i\} \]

- **PR\(_i\) : i^{th} Learner Profile**
  - General student information such as Identification data, Demographic information, etc.

- **LK\(_i\) : i^{th} Learner’s Knowledge Component**
  - Sequences of weighted visited learning objects i.e. vectors of visited learning objects or curriculum elements in which the student was interested (learner’s knowledge)

- **LEP\(_i\) : i^{th} Learner’s Educational Component**
  - Learner’s educational preferences and Learning style.
Learner’s Knowledge Component

Let $LO$ be a set of $n$ unique visited learning objects: $LO = \{LO_1, LO_2, LO_3, \ldots, LO_n\}$.

The learner knowledge model $LK_i$ corresponding to the learner $i$ is represented by a set of $p$ sessions $S^i_j$ extracted from tracked data:

$LK_i = \{S^i_1, S^i_2, S^i_3, \ldots, S^i_p\}$, where each $S^i_j$ is a subset of $k$ weighted visited learning objects $LO_l, S^i_j : < w(LO^i_1 S^i_j), w(LO^i_2 S^i_j), w(LO^i_3 S^i_j), \ldots, w(LO^i_k S^i_j) >$, where each $LO^i_k S^i_j = LO_l$ for some $l \in \{1, \ldots, n\}$, and $w(LO^i_k S^i_j)$ is the weight associated with the visited learning object $LO^i_k S^i_j$ in the session $S^i_j$ corresponding to the $i^{th}$ learner.

The learning object weight $w(LO^i_k S^i_j)$, also called, learning object interest measure $LOIM$, is computed as a function of a number of features (based on the frequency of occurrence of the learning object in a session and/or the time a learner spends when visiting a learning object).
Learner’s Knowledge Component

To compute the weights of visited learning objects within the learner’s sessions, we apply two heuristics: \textit{duration and frequency of visit}. Hence, the learning object weight \( w(LO_k^{s_j^i}) \), is obtained as follows:

\[
    w(LO_k^{s_j^i}) = \left( \frac{F(LO_k^{s_j^i})}{\sum_k F(LO_k^{s_j^i})} + \frac{T(LO_k^{s_j^i})}{\sum_k T(LO_k^{s_j^i})} \right) / 2
\]

\( F(LO_k^{s_j^i}) \) represents the frequency of visit of the learning object \( LO_k^{s_j^i} \) within the session \( S_j^i \), \( \sum_k F(LO_k^{s_j^i}) \) represents the sum of all visit frequency values related to the entire set of learning objects \( LO_k^{s_j^i} \) within the session \( S_j^i \), \( T(LO_k^{s_j^i}) \) represents the duration of a single visit to the learning object \( LO_k^{s_j^i} \) in the session \( S_j^i \), and \( \sum_k T(LO_k^{s_j^i}) \) represents the sum of all duration visits to the entire set of learning objects \( LO_k^{s_j^i} \) in the session \( S_j^i \).
Learner’s Knowledge Component

The learner’s knowledge component \( LK_i \) can be represented as a matrix \( M(p, n) \), where \( p \) is the total number of learner’s sessions and \( n \) the cardinality of unique visited learning objects.

\[
LK_i = \begin{pmatrix}
S_{1}^i & S_{2}^i & S_{3}^i & \cdots & S_{p}^i \\
\end{pmatrix}
\]

\[
m_{xy} = w(LO_{x}^{S_{y}^i}), x \in \{1, \ldots, p\}, y \in \{1, \ldots, n\}
\]

\[
= \begin{pmatrix}
w(LO_{1}^{S_{1}^i}) & \cdots & w(LO_{n}^{S_{1}^i}) \\
\vdots & \ddots & \vdots \\
w(LO_{1}^{S_{p}^i}) & \cdots & w(LO_{n}^{S_{p}^i})
\end{pmatrix}
\]
Learner’s Educational Component

Composed of the learner’s preferences among visited learning objects and his/her learning style.

Detection of the learner’s preferences:
What kind of learning object does a learner prefer?

Learning objects available in LMS are characterized by many attributes (e.g. format), each of which may have several values (e.g. for format: text, image, video, etc.) that could be preferred or not by the learner. The preferences of a learner $i$ upon these values can be represented, as a vector of interest measures: $V_{EP}^i: \{LOIM_{Att_k-LO^i_t} \}_{k,t}$

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Related values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type_LO (Learning object type)</td>
<td>{Resource, Activity}</td>
</tr>
<tr>
<td>Shape_LO (Learning object shape, if Type_LO = Activity)</td>
<td>{Exercise, Simulation, Questionnaire, Assessment, Forum, Chat, Wiki, Assignment}</td>
</tr>
<tr>
<td>Format_LO (Learning object format, if Format_LO = Resource)</td>
<td>{Text, HTML, Image, Sound, Video}</td>
</tr>
</tbody>
</table>
Let $LOIM_{Type\_LO_t^i}$, $LOIM_{Format\_LO_t^i}$, and $LOIM_{Shape\_LO_t^i}$ be the interest measures composing the vector $V_{EP}^i$, which are computed over all learner’s sessions $S^i$ as follows:

$$LOIM_{Type\_LO_t^i} = \frac{\sum_j LOIM_{Type\_LO_t^{s_j^i}}}{\text{Card}(S^i)} , \ t \in \{1, 2\}$$

$$LOIM_{Shape\_LO_t^i} = \frac{\sum_j LOIM_{Shape\_LO_t^{S_j^i}}}{\text{Card}(S^i)} , \ t \in \{1, 8\}$$

$$LOIM_{Format\_LO_t^i} = \frac{\sum_j LOIM_{Shape\_LO_t^{S_j^i}}}{\text{Card}(S^i)} , \ t \in \{1, 5\}$$

Where $\text{Card}(S^i)$ represents the number of sessions $S_j^i$ in $S^i$ belonging to the learner $i$, $LOIM_{Type\_LO_t^{S_j^i}}$, $LOIM_{Format\_LO_t^{S_j^i}}$, and $LOIM_{Shape\_LO_t^{S_j^i}}$ represent the interest measures computed over each learner’s session $S_j^i$. 
### Learner’s Educational Component

<table>
<thead>
<tr>
<th>Attributes and corresponding values</th>
<th>Interest measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type_LO</strong></td>
<td></td>
</tr>
<tr>
<td>Resource</td>
<td>$LOIM_{Type_LO}^j_{Resource}$</td>
</tr>
<tr>
<td>Activity</td>
<td>$LOIM_{Type_LO}^j_{Activity}$</td>
</tr>
<tr>
<td><strong>Shape_LO</strong></td>
<td></td>
</tr>
<tr>
<td>Exercice</td>
<td>$LOIM_{Shape_LO}^j_{Exercice}$</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>$LOIM_{Shape_LO}^j_{Questionnaire}$</td>
</tr>
<tr>
<td>test</td>
<td>$LOIM_{Shape_LO}^j_{Test}$</td>
</tr>
<tr>
<td>Forum</td>
<td>$LOIM_{Shape_LO}^j_{Forum}$</td>
</tr>
<tr>
<td>Chat</td>
<td>$LOIM_{Shape_LO}^j_{Chat}$</td>
</tr>
<tr>
<td>Wiki</td>
<td>$LOIM_{Shape_LO}^j_{Wiki}$</td>
</tr>
<tr>
<td>Navigation</td>
<td>$LOIM_{Shape_LO}^j_{Navigation}$</td>
</tr>
<tr>
<td>example</td>
<td>$LOIM_{Shape_LO}^j_{Example}$</td>
</tr>
<tr>
<td><strong>Format_LO</strong></td>
<td></td>
</tr>
<tr>
<td>Text</td>
<td>$LOIM_{Format_LO}^j_{Text}$</td>
</tr>
<tr>
<td>Html</td>
<td>$LOIM_{Format_LO}^j_{Html}$</td>
</tr>
<tr>
<td>Image</td>
<td>$LOIM_{Format_LO}^j_{Image}$</td>
</tr>
<tr>
<td>Audio</td>
<td>$LOIM_{Format_LO}^j_{Audio}$</td>
</tr>
<tr>
<td>Vidéo</td>
<td>$LOIM_{Format_LO}^j_{Video}$</td>
</tr>
</tbody>
</table>

$$LOIM_{Type_LO}^j_{t} = \frac{\sum_k w(LO_{k,t}^j)}{\text{Nbre}_{LO_t^j}}, t \in \{1, 2\}$$

$$LOIM_{Shape_LO}^j_{t} = \frac{\sum_j \text{LOIM}_{Type_LO}^j_{t}}{\text{Card}(S^t)}, t \in \{1, 2\}$$

$$LOIM_{Shape_LO}^j_{t} = \frac{\sum_k w(LO_{k,t}^j)}{\text{Nbre}_{LO_t^j}}, t \in \{1, 8\}$$

$$LOIM_{Shape_LO}^j_{t} = \frac{\sum_j \text{LOIM}_{Shape_LO}^j_{t}}{\text{Card}(S^t)}, t \in \{1, 8\}$$

$$LOIM_{Format_LO}^j_{t} = \frac{\sum_k w(LO_{k,t}^j)}{\text{Nbre}_{LO_t^j}}, t \in \{1, 5\}$$

$$LOIM_{Format_LO}^j_{t} = \frac{\sum_j \text{LOIM}_{Format_LO}^j_{t}}{\text{Card}(S^t)}, t \in \{1, 5\}$$
Learner’s Educational Component

→ Detection of the learning style (Graf et al., 2008)

FSLSM

Commonly incorporated features in LMS

- Content
- Outline
- Example
- Self-Assessment
- Exercise
- Forum
- Navigation

**Active/Reflective**
- Content_visit(-)
- Content_stay(-)
- Example_visit(+)
- Example_stay(+)
- Selfass_visit(+)
- Selfass_stay(+)
- Forum_visit(-)
- Forum_stay(-)
- Forum_post(-)

**Sensing/Intuitive**
- Ques_text(-)
- Ques_detail(+)
- Ques_overview(-)
- Ques_interpret(-)

**Visual/Verbal**
- Outline_stay(-)
- Ques_detail(+)

**Sequential/Global**
- Commonly incorporated features in LMS

ICALT 2013, Beijing, China 15/07/2013
Once learner models are built, we apply a hierarchical multi-level model based collaborative filtering approach on these models, in order to assign learners with common preferences and interests to the same groups, so that feedback from one learner can serve as a guideline for information delivery to the other learners within the same group.
Learning Styles

Educational Preferences

Level 1: Classification

Level 2: Clustering

Level 3: Clustering

Group vectors

W1, W2, W3, ..., Wm
The proposed approach has been implemented and an experiment was performed as part of a recommendation approach.

Recommendations are computed with respect to the learner’s clickstreams, his/her learning style and educational preferences, as well as exploiting similarities and dissimilarities among the learner models and educational content.

Moodle LMS is used to implement the proposed approach.

We used an online hybrid course (C2i) with 704 learners.

Tracked data was successfully extracted from various Moodle tables (primarily from mdl_log table).
Conclusions and Future Work

- Developed an approach for automatically modeling learners (groups) within LMS taking into account their educational preferences and learning styles.

- Proposed approach falls within the scope of building an educational automatic hybrid recommender system providing suitable recommendations to learners for personalized technology enhanced learning.

Future work:
- Evaluation of the recommender system.