Interactive Unknowns Recommendation in E-Learning Systems
Shan-Yun Teng1, Jundong Li2, Lo Pang-Yun Ting1, Kun-Ta Chuang1, Huan Liu2
1Dept. of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan
2School of Computing, Informatics and Decision Systems Engineering, Arizona State University, Tempe, USA

INTRODUCTION

E-learning systems should provide supplementary resources to enrich the knowledge of users' personal unknowns. Unfortunately, according to our analysis on the real log of user learning process, most users are not aware of his/her personal unknowns. Solving questions which users already understood cannot effectively enrich their knowledge because of the lack of unknowns exploration. There is no teacher for users to interact with, and users’ knowledge changes over time, which cannot easily be aware by themselves. Therefore, how to interactively exploring users’ unknowns will be the key to the success of learning new knowledge.

PROBLEM DEFINITION

Problem Statement: Let \( U = \{u_1, u_2, \ldots, u_n\} \) be the set of \( n \) users, \( Q = \{q_1, q_2, \ldots, q_m\} \) be the set of \( m \) questions, and \( C = \{c_1, c_2, \ldots, c_k\} \) be the set of \( k \) concepts. For a user-question rating matrix \( R \in \mathbb{R}^{n \times m} \), \( R_{ij} = 1 \) if user \( u_i \) correctly answers question \( q_j \), \( R_{ij} = -1 \) otherwise.

Given: users \( U \), questions \( Q \), concepts \( C \), user-question rating matrix \( R \), and the proposed graphs \( G_{uc}, G_{cr}, \) and \( G_{cq} \).

Select: a set of \( k \) questions (user unknowns) for users by interactively exploring rating matrix \( R \) along with proposed graph structures encoded in the adjacency matrices \( A^u \), \( A^r \), and \( A^q \).

CagMab Framework

In this paper, we propose the CagMab framework incorporating concept-aware graph embedding into the multi-armed bandit. We propose concept-aware graph embeddings to preserve the graph structures into matrix factorization (MF) model, which helps to capture the interactions between question-concept, concept closeness and concept prerequisite. Then, we are able to well optimize the latent factors of users and questions.

\[
\begin{align*}
\min_{\beta, \gamma} \beta(\Omega_{c_1} + \gamma(\Omega_{c_1} + \Omega_{c_2} + \Omega_{c_3})) &= \frac{1}{2} \sum_{i,j} \Omega_{ij} (R_{ij} - \beta \cdot q_i) \cdot (q_j) + \frac{\lambda}{2} \|U\|^2 + \frac{\lambda}{2} \|Q\|^2 \\
&- \gamma \left( \sum_{i,j} w_{ij} \log(p(c_i|c_j) + \log(p(q_i|c_j))) + \sum_{c_i \in C, c_j \in N_2(c_i)} \sum_{q_k \in E} r_{ij}(c_i, c_j) \log p(q_k|c_i) \right) \\
&+ \sum_{c_i \in E} r_{ij}(c_i, c_j) \log p(q_k|c_i)
\end{align*}
\]

(1)

We introduce both the latent factors and contextual factors of users and questions to the reward with user affinity graph in a multi-armed bandit framework.

\[
r_{ui, qj} = \mathbb{E}[r_{ui, qj}|\bar{x}_{ui}, \bar{a}_u, \bar{u}, \bar{q}j] = \bar{x}^T_{ui, qj} \bar{a} + \bar{u} \bar{q}j + \epsilon_z.
\]

(2)

INTERACTIVE RECOMMENDATION PERFORMANCE

<table>
<thead>
<tr>
<th>Metric</th>
<th>Native</th>
<th>TS</th>
<th>UCB</th>
<th>LUCB</th>
<th>COLB</th>
<th>CAGMAB</th>
<th>CAGMAB+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision@5</td>
<td>0.1400</td>
<td>0.1320</td>
<td>0.1320</td>
<td>0.1300</td>
<td>0.1400</td>
<td>0.1400</td>
<td>0.1400</td>
</tr>
<tr>
<td>Recall@5</td>
<td>0.0404</td>
<td>0.1100</td>
<td>0.1160</td>
<td>0.1170</td>
<td>0.1255</td>
<td>0.1099</td>
<td>0.1099</td>
</tr>
<tr>
<td>Precision@10</td>
<td>0.1400</td>
<td>0.1440</td>
<td>0.1410</td>
<td>0.1440</td>
<td>0.1319</td>
<td>0.1340</td>
<td>0.1340</td>
</tr>
<tr>
<td>Recall@10</td>
<td>0.2222</td>
<td>0.2396</td>
<td>0.2000</td>
<td>0.2512</td>
<td>0.2256</td>
<td>0.2114</td>
<td>0.2178</td>
</tr>
<tr>
<td>AUC@10</td>
<td>0.6640</td>
<td>0.6750</td>
<td>0.6070</td>
<td>0.6400</td>
<td>0.6400</td>
<td>0.6400</td>
<td>0.6400</td>
</tr>
<tr>
<td>AUC@10</td>
<td>0.6640</td>
<td>0.6750</td>
<td>0.6070</td>
<td>0.6400</td>
<td>0.6400</td>
<td>0.6400</td>
<td>0.6400</td>
</tr>
</tbody>
</table>

Table 1: Performance comparison in terms of Precision, Recall and AUC.

Figure 1: Performance comparison on dataset with 10% cold-start users in terms of Precision and Recall.

PARAMETER SENSITIVITY

Figure 2: Parameter sensitivity of CagMab.

CASE STUDIES

Figure 3: The richness and the diversity of recommended questions.

Figure 4: The ratio of known unknowns and unknown unknowns in recommended questions.