Modeling Retail Transaction Data for Personalized Shopping Recommendation

Pengfei Wang, Jiafeng Guo, Yanyan Lan, Xueqi Cheng
CAS-NDST Lab, Institute of Computing Technology, Chinese Academy of Sciences
wangpengfei@software.ict.ac.cn, {guojiafeng, lanyanyan, cxq}@ict.ac.cn

1. BACKGROUND AND MOTIVATION

Massive transaction data has been routinely recorded in offline retail, which convey rich preference information on brands and goods from customers. Personalized Recommendation based on these valuable transaction data is a critical task.

2. DEFINITION

Inspired by association rules, we introduce association patterns as basic units to capture the correlation between products.

3. PAP MODEL

Key idea: reduce association patterns into a low-dimensional shopping interest space, inference the interest of each individual, and then provide personalized shopping recommendations.

Probabilistic model over Association Patterns:

4. RECOMMENDATION

Inference of User Preference

Probability of the k-th shopping interest:

Personalized Recommendation

By sorting the products according to $P(i|u)$, we can recommend top-k products to the user:

5. EXPERIMENT

1. Statistics of datasets:

<table>
<thead>
<tr>
<th>Dataset</th>
<th># users</th>
<th># products</th>
<th># transactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>BeiRen</td>
<td>19515</td>
<td>1442</td>
<td>243901</td>
</tr>
<tr>
<td>Tafeng</td>
<td>7144</td>
<td>6984</td>
<td>37269</td>
</tr>
</tbody>
</table>

2. Results on BeiRen and Tafeng:

Two datasets show that our PAP model perform better than the state-of-the-art recommendation methods.

6. CONCLUSION

Contributions

- We introduce association patterns as basic units to capture the correlation between products.
- We proposed a novel Probabilistic model over the Association Patterns for personalized recommendation.
- Experiments showed that our method outperformed than state-of-the-art recommendation methods.

Future Work

- In the future we will try to explore long-term and short-term association patterns, and analyze the impact of two patterns to personalized recommendation.