Multi-Aspect Collaborative Filtering based on Linked Data for Personalized Recommendation

Han-Gyu Ko
kohangyu@kaist.ac.kr
Joo-Sik Son
sonjoosik@kaist.ac.kr
In-Young Ko
iko@kaist.ac.kr
School of Computing, Korea Advanced Institute of Science and Technology
291 Daehak-ro, Yuseong-gu, Daejeon, 305-701, Republic of Korea

Abstract
Since users often consider more than one aspect when they choose an item, relevant researches introduced multi-criteria recommender systems and showed that multi-criteria ratings add values to the existing collaborative-filtering-based recommender systems to provide more accurate recommendation results to users. However, all the previous works require multi-criteria ratings given by users explicitly while most of the existing datasets such as Netflix and MovieLens include only single-criterion ratings. Therefore, to take advantage of multi-criteria recommendation, there must be a way to extract necessary aspects and analyze users’ preferences on those aspects from the given single-criterion type of dataset. In this paper, we propose an approach of utilizing semantic information of items to extracting essential aspects for performing multi-aspect collaborative filtering to recommend users with items in a personalized manner.

Main characteristics of existing Collaborative Filtering based recommender systems
• Utilize the feedback information generated on items to predict users’ preferences
• Most of the feedback information is single criterion such as users’ ratings on items

An alternative and its limitation
• ‘Multi-criteria ratings’ allow more accurate CF-based recommendation
• It’s a burden for users to provide more than one feedback on a item!

Multi-Aspect Collaborative Filtering based on Linked Data
• Goal: overcome the limitation and take the advantage of multi-criteria CF-based recommendation
• Intuition: enrich item information semantically by associating relevant concepts from Linked Data
• Use associated concepts to measure users’ preferences on multiple aspects

Multi-Aspect Matrix Localization
• Generate groups of users and items based on their similarity for each aspect
- By using concept groups, measure the similarity between the set of users and items
- Matrix Completion
  Matrix Completion: \( M - (U', V') = \arg \min \sum_{i,j} (M_{i,j} - [U'V'^T])^2 \)
- Matrix Integration
  \( \text{EstimatedRating}(i,j) = \sum_{k} W(i,k) \times \text{Rating}(k,j) \)

Integration of Prediction Results in Multi-Aspects
• Assign weight values to each aspect differently according to each user
  - based on Analytic Hierarchical Process (AHP)
  - then integrate the prediction results for each aspect

[Comparison of Recommendation Accuracy]

<table>
<thead>
<tr>
<th>Dataset/Method</th>
<th>MAE</th>
<th>RMSE</th>
<th>Avg. Precision</th>
<th>NDCG</th>
<th>NDCG@10</th>
</tr>
</thead>
<tbody>
<tr>
<td>UserBud</td>
<td>0.6941</td>
<td>0.8911</td>
<td>0.8340</td>
<td>0.9134</td>
<td></td>
</tr>
<tr>
<td>ItemBud</td>
<td>0.6725</td>
<td>0.8583</td>
<td>0.8420</td>
<td>0.9165</td>
<td></td>
</tr>
<tr>
<td>RegSVD</td>
<td>0.6922</td>
<td>0.8686</td>
<td>0.8435</td>
<td>0.9174</td>
<td></td>
</tr>
<tr>
<td>NMF</td>
<td>0.6925</td>
<td>0.8837</td>
<td>0.8271</td>
<td>0.9069</td>
<td></td>
</tr>
<tr>
<td>PMF</td>
<td>0.8730</td>
<td>1.0494</td>
<td>0.7938</td>
<td>0.8886</td>
<td></td>
</tr>
<tr>
<td>BPMF</td>
<td>0.6562</td>
<td>0.8409</td>
<td>0.8489</td>
<td>0.9196</td>
<td></td>
</tr>
<tr>
<td>LLOMA</td>
<td>0.6941</td>
<td>0.8911</td>
<td>0.8340</td>
<td>0.9134</td>
<td></td>
</tr>
<tr>
<td>MultiAspect</td>
<td>0.5081</td>
<td>0.7182</td>
<td>0.8945</td>
<td>0.9601</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion and Future Work

Main Contributions
We proposed and developed ...
• a framework of utilizing Linked Data to expand keywords from item metadata
• a way of identifying similar interests of users by using concept groups with given aspects
• an effective way of aggregating the prediction results from sub-matrices in multiple aspects

Future Work
• Develop an optimal method to decide weight values for each user
  - based on machine learning techniques such as gradient decent
• Make the proposed approach scalable
  - matrix completion is a time-consuming process
  - use cluster machines to run the approach in a concurrent manner