STAR: croSSs-plaTform App Recommendation

Da Cao, Xiangnan He, Liqiang Nie, Xiaochi Wei, Xia Hu, Shunxiang Wu, Tat-Seng Chua

This work was finished when Da Cao was a Ph.D. student in XMU and a visiting student in NUS.
Currently Da Cao is an assistant professor in HNU.
caoda0721@gmail.com

Submit: June 2016
Accepted: Nov 2016
Published: July 2017
Outline

• Background
• Proposed Method
• Experiments and Results
• Conclusion
App Development

Mobile Network → App-Driven Life → Multi-Platform → Overwhelmed
Two App Recommendation solutions

Single Platform App Recommendation

- Smartphone
  - Training Data
  - Training
  - Specific Features
  - Recommendation

- Tablet
  - Training Data
  - Training
  - Specific Features
  - Recommendation

- Computer
  - Training Data
  - Training
  - Specific Features
  - Recommendation

Cross-Platform App Recommendation

- Smartphone
  - Training Data

- Tablet
  - Training Data

- Computer
  - Training Data

- Joint Training
  - Common Features
  - Specific Features
  - Recommendation

Who wins?

Single Platform vs Cross-Platform

Russia vs Croatia
Challenges

• Platform Variance
• Data Heterogeneity
• Data Sparsity
• Cold-Start Problem
Outline

• Background
• Proposed Method
• Experiments and Results
• Conclusion
Proposed Method

\[ \hat{r}_{ijs} = \mu + b_u(i) + b_v(js) + u_i^T v_{js} \]

\[ v_{js} = w_j + M\theta_{js} \]

\[ \hat{r}_{ijs} = \mu + b_u(i) + b_w(j) + u_i^T (w_j + M\theta_{js}) \]

\[ \mathcal{L}(\Theta) = \frac{1}{2} \sum_{(i,j,s) \in \mathcal{R}} (r_{ijs} - \hat{r}_{ijs})^2 + \lambda \|\Theta\|^2 \]

where \( \Theta = \{U, W, M, b_u, b_w\} \)
Cold-Start Problems

New-user cold-start

- **iPhone**
  - Alex
  - Social enhanced
  - App X

- **iPad**
  - Alex
  - Nice looking
  - App Y

- **iMac**
  - Alex
  - ?
  - App Z

New-App cold-start

- **iPhone**
  - Alex
  - Non-free
  - App X

- **iPad**
  - Bob
  - Easy to operate
  - App X

- **iMac**
  - Clark
  - ?
  - App X

The question mark “?” stands for the ratings that we wish to predict, and the label “new” means the user or App is new to the platform and has no rating history on it.
Outline

• Background
• Proposed Method
• Experiments and Results
• Conclusion
Dataset and Evaluation

We selected users who rated at least once on both of these platforms.

### Table I. Some Statistics of the iphone-iPad Dataset

<table>
<thead>
<tr>
<th></th>
<th>Amount</th>
<th>Min. #ratings</th>
<th>Max. #ratings</th>
<th>Avg. #ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td>112,031</td>
<td>2</td>
<td>219</td>
<td>2.86</td>
</tr>
<tr>
<td>App-iPhone</td>
<td>2,704</td>
<td>1</td>
<td>15,946</td>
<td>62.31</td>
</tr>
<tr>
<td>App-iPad</td>
<td>2,704</td>
<td>1</td>
<td>11,846</td>
<td>56.23</td>
</tr>
</tbody>
</table>

We selected users who had at least two ratings on all platforms.

### Table II. Some Statistics of the iphone-iPad-iMac Dataset

<table>
<thead>
<tr>
<th></th>
<th>Amount</th>
<th>Min. #ratings</th>
<th>Max. #ratings</th>
<th>Avg. #ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>User</td>
<td>121,905</td>
<td>2</td>
<td>21</td>
<td>2.21</td>
</tr>
<tr>
<td>App-iPhone</td>
<td>201</td>
<td>1</td>
<td>64,482</td>
<td>1,117.37</td>
</tr>
<tr>
<td>App-iPad</td>
<td>201</td>
<td>1</td>
<td>5,572</td>
<td>206.62</td>
</tr>
<tr>
<td>App-iMac</td>
<td>201</td>
<td>1</td>
<td>439</td>
<td>13.97</td>
</tr>
</tbody>
</table>

Rating Prediction

- MAE
- RMSE

Top-N Recommendation

- Recall
- NDCG

7/16/2018
Research Questions

• *(RQ1).* How does STAR perform as compared to other state-of-the-art competitors?

• *(RQ2).* How is the performance of STAR in handling the new-user and new-App cold-start problems?

• *(RQ3).* Whether the rated App on current platform is the user's preferable one as compared to the same App on other unrated platforms?

• *(RQ4).* How do the common features and specific features of Apps contribute to the overall effectiveness of STAR?

• *(RQ5).* In addition to rating prediction that is prevalent to a recommendation algorithm, how does STAR perform in the more practical top-N recommendation?
Baseline Methods

• SVD++ [Koren 2008] (Collaborative Filtering)
• RMR [Ling et al. 2014] (Semantics Enhanced Recommendation)
• CTR [Wang and Blei 2011] (Semantics Enhanced Recommendation)
• FM [Rendle et al. 2011] (Context-Aware Recommender System)
• CMF [Singh and Gordon 2008] (Cross-Domain Recommender System)
• WMF [Hu et al. 2008] (Collaborative Filtering)
• Popular (Non-personalized method)
Overall Performance Comparisons (RQ1)

Table III. Performance comparison of various methods on the iPhone-iPad dataset regarding RMSE and MAE.

<table>
<thead>
<tr>
<th>Methods</th>
<th>MAE</th>
<th>RMSE</th>
<th>p-value (MAE)</th>
<th>p-value (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVD++</td>
<td>0.7823 ± 0.003</td>
<td>1.1218 ± 0.004</td>
<td>2.44e-10</td>
<td>1.02e-10</td>
</tr>
<tr>
<td>RMR</td>
<td>0.7679 ± 0.002</td>
<td>1.0716 ± 0.003</td>
<td>5.83e-09</td>
<td>7.16e-08</td>
</tr>
<tr>
<td>FM</td>
<td>0.8002 ± 0.003</td>
<td>1.0887 ± 0.003</td>
<td>3.05e-11</td>
<td>2.06e-10</td>
</tr>
<tr>
<td>CMF</td>
<td>0.7701 ± 0.003</td>
<td>1.0977 ± 0.003</td>
<td>2.96e-09</td>
<td>7.21e-10</td>
</tr>
<tr>
<td>STAR</td>
<td>0.7560 ± 0.003</td>
<td>1.0595 ± 0.003</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table IV. Performance comparison of various methods on the iPhone-iPad-iMac dataset regarding RMSE and MAE.

<table>
<thead>
<tr>
<th>Methods</th>
<th>MAE</th>
<th>RMSE</th>
<th>p-value (MAE)</th>
<th>p-value (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVD++</td>
<td>0.6040 ± 0.002</td>
<td>0.9389 ± 0.002</td>
<td>2.88e-09</td>
<td>5.79e-11</td>
</tr>
<tr>
<td>RMR</td>
<td>0.5985 ± 0.003</td>
<td>0.9333 ± 0.004</td>
<td>1.77e-08</td>
<td>7.61e-07</td>
</tr>
<tr>
<td>FM</td>
<td>0.6117 ± 0.003</td>
<td>0.9352 ± 0.003</td>
<td>5.55e-10</td>
<td>2.19e-08</td>
</tr>
<tr>
<td>CMF</td>
<td>0.6035 ± 0.003</td>
<td>0.9371 ± 0.003</td>
<td>3.30e-09</td>
<td>6.60e-10</td>
</tr>
<tr>
<td>STAR</td>
<td>0.5889 ± 0.002</td>
<td>0.9261 ± 0.003</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Handling Cold-Start Problems (RQ2)

Table V. Performance comparison of various methods in handling new-user cold-start problem.

<table>
<thead>
<tr>
<th>Methods</th>
<th>MAE</th>
<th>RMSE</th>
<th>p-value (MAE)</th>
<th>p-value (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMR</td>
<td>0.7003 ± 0.003</td>
<td>1.0592 ± 0.002</td>
<td>2.67e-09</td>
<td>3.97e-10</td>
</tr>
<tr>
<td>FM</td>
<td>0.7223 ± 0.003</td>
<td>1.0690 ± 0.003</td>
<td>7.67e-11</td>
<td>1.05e-10</td>
</tr>
<tr>
<td>CMF</td>
<td>0.7041 ± 0.004</td>
<td>1.0783 ± 0.003</td>
<td>1.10e-09</td>
<td>4.04e-11</td>
</tr>
<tr>
<td>STAR</td>
<td>0.6849 ± 0.003</td>
<td>1.0344 ± 0.003</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Table VI. Performance comparison of various methods in handling new-App cold-start problem.

<table>
<thead>
<tr>
<th>Methods</th>
<th>MAE</th>
<th>RMSE</th>
<th>p-value (MAE)</th>
<th>p-value (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTR</td>
<td>0.8746 ± 0.003</td>
<td>1.2310 ± 0.003</td>
<td>6.76e-10</td>
<td>6.98e-10</td>
</tr>
<tr>
<td>FM</td>
<td>0.9017 ± 0.003</td>
<td>1.2452 ± 0.004</td>
<td>2.64e-11</td>
<td>1.76e-11</td>
</tr>
<tr>
<td>CMF</td>
<td>0.8736 ± 0.002</td>
<td>1.2471 ± 0.002</td>
<td>8.17e-10</td>
<td>2.48e-11</td>
</tr>
<tr>
<td>STAR</td>
<td>0.8529 ± 0.002</td>
<td>1.2063 ± 0.003</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
User Preference on App-Platform (RQ3)

1. Our method favors the current platform better.
2. The gap between rating predictions of current platform and other platforms on the iPhone-iPad dataset is larger than that of the iPhone-iPad-iMac dataset.
Justification of Common Features and Specific Features (RQ4)

Table VII. Importance comparisons of common features and specific features on the iPhone-iPad dataset.

<table>
<thead>
<tr>
<th>Methods</th>
<th>MAE</th>
<th>RMSE</th>
<th>p-value (MAE)</th>
<th>p-value (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cSTAR</td>
<td>$0.7771 \pm 0.004$</td>
<td>$1.0824 \pm 0.003$</td>
<td>$7.57e-10$</td>
<td>$8.72e-09$</td>
</tr>
<tr>
<td>sSTAR</td>
<td>$0.7842 \pm 0.003$</td>
<td>$1.1042 \pm 0.002$</td>
<td>$2.37e-10$</td>
<td>$5.39e-13$</td>
</tr>
<tr>
<td>STAR</td>
<td>$0.7560 \pm 0.003$</td>
<td>$1.0595 \pm 0.003$</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Table VIII. Importance comparisons of common features and specific features on the iPhone-iPad-iMac dataset.

<table>
<thead>
<tr>
<th>Methods</th>
<th>MAE</th>
<th>RMSE</th>
<th>p-value (MAE)</th>
<th>p-value (RMSE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cSTAR</td>
<td>$0.6092 \pm 0.003$</td>
<td>$0.9364 \pm 0.003$</td>
<td>$8.33e-10$</td>
<td>$2.13e-07$</td>
</tr>
<tr>
<td>sSTAR</td>
<td>$0.6299 \pm 0.003$</td>
<td>$0.9471 \pm 0.002$</td>
<td>$5.31e-11$</td>
<td>$1.24e-09$</td>
</tr>
<tr>
<td>STAR</td>
<td>$0.5889 \pm 0.002$</td>
<td>$0.9261 \pm 0.003$</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
Evaluation of Top-N Recommendation (RQ5)

Table IX. Evaluation of Top-N Recommendation on the iphone-iPad Dataset

<table>
<thead>
<tr>
<th>Methods</th>
<th>Recall@100</th>
<th>NDCG@100</th>
<th>p-value(Recall)</th>
<th>p-value(NDCG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Popular</td>
<td>0.5227 ± 0.000</td>
<td>0.1757 ± 0.000</td>
<td>1.58e-11</td>
<td>1.48e-11</td>
</tr>
<tr>
<td>WMF</td>
<td>0.5695 ± 0.003</td>
<td>0.2158 ± 0.002</td>
<td>1.02e-10</td>
<td>2.12e-09</td>
</tr>
<tr>
<td>CMF-WMF</td>
<td>0.5696 ± 0.002</td>
<td>0.2241 ± 0.003</td>
<td>2.74e-08</td>
<td>3.66e-08</td>
</tr>
<tr>
<td>STAR-WMF</td>
<td><strong>0.5782 ± 0.002</strong></td>
<td><strong>0.2321 ± 0.003</strong></td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Table X. Evaluation of Top-N Recommendation on the iphone-iPad-iMac Dataset

<table>
<thead>
<tr>
<th>Methods</th>
<th>Recall@10</th>
<th>NDCG@10</th>
<th>p-value(Recall)</th>
<th>p-value(NDCG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Popular</td>
<td>0.6133 ± 0.000</td>
<td>0.5005 ± 0.000</td>
<td>6.21e-12</td>
<td>4.42e-11</td>
</tr>
<tr>
<td>WMF</td>
<td>0.6686 ± 0.002</td>
<td>0.5321 ± 0.003</td>
<td>3.12e-09</td>
<td>9.19e-09</td>
</tr>
<tr>
<td>CMF-WMF</td>
<td>0.6745 ± 0.003</td>
<td>0.5379 ± 0.004</td>
<td>2.38e-08</td>
<td>1.63e-07</td>
</tr>
<tr>
<td>STAR-WMF</td>
<td><strong>0.6834 ± 0.003</strong></td>
<td><strong>0.5434 ± 0.003</strong></td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
Outline

• Background
• Proposed Method
• Experiments and Results
• Conclusion
Challenges Solved

• Platform Variance
• Data Heterogeneity
• Data Sparsity
• Cold-Start Problem
Website

Data & code are available at http://apprec.wixsite.com/star

Cross-Platform App Recommendation by Jointly Modeling Ratings and Texts

Introduction

Over the last decade, the renaissance of Web technologies has transformed the online world into an application (App) driven society. While the abundant Apps have provided great convenience, their sheer number also leads to severe information overload, making it difficult for users to identify desired Apps. To alleviate the information overloading issue, recommender systems have been proposed and deployed for the online App domain. However, existing work on App recommendation has largely focused on one single platform (e.g., smartphones), while ignores the rich and relevant data of other platforms (e.g., tablets and computers).
Thanksgiving

Co-authors:

Xiangnan He (NUS)
Liqiang Nie (SDU)
Xiaochi Wei (BIT)
Xia Hu (Texas A&M)
Shunxiang Wu (XMU)
Tat-Seng Chua (NUS)

And all audiences...