Transfer Learning from APP Domain to News Domain for Dual Cold-Start Recommendation

Jixiong Liu\textsuperscript{1}, Jiakun Shi\textsuperscript{1}, Wanling Cai\textsuperscript{1}, Bo Liu\textsuperscript{2}, Weike Pan\textsuperscript{1}
Qiang Yang\textsuperscript{2}\textsuperscript{*} and Zhong Ming\textsuperscript{1}\textsuperscript{*}

\{1455606137,1033150729,382970614\}@qq.com, \{bliuab,qyang\}@cse.ust.hk, \{panweike,mingz\}@szu.edu.cn

\textsuperscript{1}College of Computer Science and Software Engineering
Shenzhen University, Shenzhen, China

\textsuperscript{2}Department of Computer Science and Engineering
Hong Kong University of Science and Technology, Hong Kong, China
Problem Definition (1/3)

In our studied problem, we have an APP domain and a news domain:

- In the APP domain, we have a set of triples, i.e., \((u, g, G_{ug})\), denoting that user \(u\) has installed \(G_{ug}\) times of mobile apps belonging to the genre \(g\).

- In the news domain, we have a set of quadruples, i.e., \((u, i, c_1(i), c_2(i))\), denoting that user \(u\) has read an item \(i\) belonging to a level-1 category \(c_1(i)\) and a level-2 category \(c_2(i)\).

Note that we only make use of items’ category information, but not content information.
Our goal is to recommend a ranked list of latest news articles (i.e., \textit{new items}) to each \textit{new user} who has not read any news articles before.

- It is a \textit{new user cold-start} and \textit{new item cold-start} problem, which is thus termed as \textit{dual cold-start recommendation} (DCSR).
Figure: An illustration of *neighborhood-based transfer learning* (NTL) for *dual cold-start recommendation* (DCSR).

Note that each entry in the user-category matrix $\mathbf{C}$ denotes the number of items belonging to a certain category that a user has read.
Challenges

- **New user cold-start challenge**, i.e., the target users (to whom we will provide recommendations) have not read any items before.
- **New item cold-start challenge**, i.e., the target items (that we will recommend to the target users) are totally new for all users.

Most existing recommendation algorithms are not applicable.
Overall of Our Solution

*Neighborhood-based Transfer Learning (NTL)*

- For the **new user cold-start challenge**: we transfer the knowledge of neighborhood of the cold-start users from an APP domain to a news domain.
- For the **new item cold-start challenge**: we design a category-level preference to replace the traditional item-level preference because the latter is not applicable for the new items.
Advantage of Our Solution

- NTL is able to make use of the users’ app installation behaviors for news recommendation in a simple but effective way
Notations

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u$</td>
<td>user id</td>
</tr>
<tr>
<td>$i$</td>
<td>item id</td>
</tr>
<tr>
<td>$g$</td>
<td>genre id</td>
</tr>
<tr>
<td>$C_1$</td>
<td>a set of level-1 categories, $c_1 \in C_1$</td>
</tr>
<tr>
<td>$C_2$</td>
<td>a set of level-2 categories, $c_2 \in C_2$</td>
</tr>
<tr>
<td>$N_{u, c_1}$</td>
<td>the number of read items (by user $u$) belonging to a level-1 category $c_1$</td>
</tr>
<tr>
<td>$N_{u, c_2}$</td>
<td>the number of read items (by user $u$) belonging to a level-2 category $c_2$</td>
</tr>
<tr>
<td>$N_{c_1}$</td>
<td>the number of read items (by all users) belonging to a level-1 category $c_1$</td>
</tr>
<tr>
<td>$N_{c_2}$</td>
<td>the number of read items (by all users) belonging to a level-2 category $c_2$</td>
</tr>
<tr>
<td>$p_{c_1}$</td>
<td>the popularity of the level-1 category $c_1$ among the users</td>
</tr>
<tr>
<td>$p_{c_2}$</td>
<td>the popularity of the level-2 category $c_2$ among the users</td>
</tr>
<tr>
<td>$N_u$</td>
<td>a set of neighbors of user $u$</td>
</tr>
</tbody>
</table>
Cross-Domain Preference Assumption

- Users with similar app-installation behaviors are likely to have similar tastes in news articles
  - For instance, two users with the installed apps of the same genre *business* may both prefer news articles on topics like *finance*. 
Cosine Similarity

With the cross-domain preference assumption, we first calculate the cosine similarity between a cold-start user $u$ and a warm-start user $u'$ in the APP domain as follows,

$$s_{u,u'} = \frac{G_u \cdot G^T_{u'}}{\sqrt{G_u \cdot G^T_u} \cdot \sqrt{G_{u'} \cdot G^T_{u'}}},$$  \hspace{1cm} (1)$$

where $G_u$ is a row vector w.r.t. user $u$ from the user-genre matrix $G$. 
Once we have calculated the cosine similarity, for each cold-start user $u$, we first remove users with a small similarity value (e.g., $s_{u,u'} < 0.1$), and then take some (e.g., 100) most similar users to construct a neighborhood $\mathcal{N}_u$. 
The item-level preference prediction rule for user $u$ to item $i$ is as follows,

$$\hat{r}_{u,i} = \frac{1}{|\mathcal{N}_u|} \sum_{u' \in \mathcal{N}_u} \hat{r}_{u',i},$$

(2)

where $\mathcal{N}_u$ is a set of nearest neighbors of user $u$ in terms of a certain similarity measurement such as cosine similarity, and $\hat{r}_{u',i}$ is the estimated preference of user $u'$ (a close neighbor of user $u$) to item $i$.

For the item-level preference $\hat{r}_{u',i}$ in Eq.(2), we are not able to have such a score directly because the item $i$ is new for all users, including the warm-start users and the target cold-start user $u'$. 
Category-Level Preference

We thus propose to approximate the item-level preference using a category-level preference,

\[ \hat{r}_{u',i} \approx \hat{r}_{u',c(i)}, \]

(3)

where \( c(i) \) can be the level-1 category or level-2 category. We then have two types of category-level preferences,

\[ \hat{r}_{u',c(i)} = \hat{r}_{u',c_1(i)} = N_{u',c_1(i)}, \]

(4)

\[ \hat{r}_{u',c(i)} = \hat{r}_{u',c_2(i)} = N_{u',c_2(i)}, \]

(5)

where \( N_{u',c_1(i)} \) and \( N_{u',c_2(i)} \) denote the number of read items (by user \( u' \)) belonging to the level-1 category \( c_1(i) \) and the level-2 category \( c_2(i) \), respectively.
Finally, with the Eqs.(3-5), we can rewrite Eq.(2) as follows,

\[
\hat{r}_{u,i} \approx \frac{1}{|N_u|} \sum_{u' \in N_u} N_{u',c_1(i)}, \\
\hat{r}_{u,i} \approx \frac{1}{|N_u|} \sum_{u' \in N_u} N_{u',c_2(i)},
\]

which will be used for preference prediction in our empirical studies.

Specifically, the neighborhood $N_u$ addresses the **new user cold-start challenge**, and the category-level preference $N_{u',c_1(i)}$ or $N_{u',c_2(i)}$ addresses the **new item cold-start challenge**.
In the APP domain, we have 827,949 users and 53 description terms (i.e., genres) of the users’ installed mobile apps, where the genres are from Google Play.

Considering our target task of news recommendation, we removed 14 undiscriminating or irrelevant genres such as tools, communication, social, entertainment, productivity, weather, dating, etc.

Finally, we have a matrix $G$ with 827,949 users (or rows) and 39 genres (or columns), where each entry represents the number of times that a user has installed apps belonging to a genre.
Datasets (2/2)

In the news domain, we have two sets of data, including a training data and a test data.

The training data spans from 10 January 2017 to 30 January 2017, and contains 806,167 users, 747,643 items (i.e., news articles), and 16,199,385 unique (user, item) pairs.

The test data are from the data on 31 January 2017, which contains 3,597 new users, 28,504 new items (i.e., news articles), and 4,813 unique (user, item) pairs.

Note that we have $|C_1| = 26$ level-1 categories and $|C_2| = 222$ level-2 categories about the items in the news domain.
Evaluation Metrics

- precision@15
- recall@15
- F1@15
- NDCG@15
- 1-call@15
Baselines

- Random recommendation (Random): we randomly select $K = 15$ items in the test data for each cold-start user.

- Popularity-based ranking via level-1 category (PopRank-C1): we use $\hat{r}_i = p_{c_1(i)}$ for preference prediction.

- Popularity-based ranking via level-2 category (PopRank-C2): we use $\hat{r}_i = p_{c_2(i)}$ for preference prediction.

In PopRank-C1 (or PopRank-C2), for the most popular level-1 (or level-2) category, there may be more than $K = 15$ items (i.e., articles) in the test data, we then randomly take $K$ items (i.e., articles) from that level-1 (or level-2) category for recommendation.
Experiments

Results (1/3)

<table>
<thead>
<tr>
<th>Method</th>
<th>Prec@15</th>
<th>Rec@15</th>
<th>F1@15</th>
<th>NDCG@15</th>
<th>1-call@15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>5.56E-05</td>
<td>5.84E-04</td>
<td>9.78E-05</td>
<td>2.27E-04</td>
<td>8.34E-04</td>
</tr>
<tr>
<td>PopRank-C1</td>
<td>5.00E-05</td>
<td>5.59E-04</td>
<td>9.02E-05</td>
<td>2.38E-04</td>
<td>7.51E-04</td>
</tr>
<tr>
<td>PopRank-C2</td>
<td>1.46E-04</td>
<td>1.74E-03</td>
<td>2.65E-04</td>
<td>6.48E-04</td>
<td>2.20E-03</td>
</tr>
<tr>
<td>NTL-C1</td>
<td>0.0053</td>
<td>0.0645</td>
<td>0.0095</td>
<td>0.0255</td>
<td>0.0734</td>
</tr>
<tr>
<td>NTL-C2</td>
<td>0.0040</td>
<td>0.0501</td>
<td>0.0073</td>
<td>0.0206</td>
<td>0.0567</td>
</tr>
</tbody>
</table>

Observations:

- The overall performance shows the effectiveness of our proposed neighborhood-based transfer learning solution.

- The performance of PopRank-C2 and PopRank-C1 are rather poor because they are non-personalized methods.

- NTL-C1 performs better as expected because the level-1 category may introduce more smoothing effect for the cold-start problem.
Experiments

Results (2/3)

Figure: Recommendation performance of our NTL with level-1 category (NTL-C1) using different neighborhood sizes.

Observation:

- The results are relatively stable with different numbers of neighbors, and configuring it as 100 usually produces the best performance.
Figure: Recommendation performance of our NTL with level-1 category (NTL-C1) using random neighborhood and transferred neighborhood.

Observation:
- The neighborhood constructed using the app-installation behaviors is better than that of the random counterpart, which shows that the two domains are related and can indeed transfer knowledge from one domain to the other.
Conclusion

- We study an important and challenging news recommendation problem called *dual cold-start recommendation* (DCSR).

- We propose a *neighborhood-based transfer learning* (NTL) solution, which is able to address the new user cold-start challenge and the new item cold-start challenge by the transferred neighborhood from the APP domain and the category-level preferences in the news domain, respectively.

- For future works, we are interested in selecting some representative genres and categories in two domains and building a mapping between them, which will be further used to study the neighborhood of the items.
Thank you!

- We thank the anonymous reviewers for their expert comments and constructive suggestions.
- We thank the support of Natural Science Foundation of China (NSFC) Nos. 61502307 and 61672358, China National 973 project No. 2014CB340304, and Hong Kong CERG projects Nos. 16211214, 16209715 and 16244616.