Hete-CF: Social-Based Collaborative Filtering Recommendation using Heterogeneous Relations

Chen Luo\textsuperscript{1}, Wei Pang\textsuperscript{2}, Zhe Wang\textsuperscript{1}, Chenghua Lin\textsuperscript{2}

\textsuperscript{1}Jilin University, China
\textsuperscript{2}Aberdeen University, UK

http://www.chen-loo.com/
Outline

• Background and Motivation

• The Hete-CF Model

• Experimental Result

• Conclusion
CF-based Recommendation

• Use user-item rating matrix $\mathbf{R}$ to make recommendation.

$$\min_{\mathbf{U}, \mathbf{V}} \sum_{i=0}^{m} \sum_{j=0}^{n} (U_i^T V_j - R_{i,j})^2 + \lambda (\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2)$$

• $\mathbf{U}$ and $\mathbf{V}$ are lower rank representations of users and items.

• Then the ratings can be calculated as

$$\hat{R}_{i,j} = U_i^T V_j$$

• Suffered by Data Sparsity and Cold start problems.
Social Recommendation

• Using the information from social networks to help recommendation.
  • Trust-MF (IJCAI 2013) Uses trust relationship between users to improve recommendation
  • Hete-MF (IJCAI-HINA 2013) Uses item (Entity) similarity information to improve recommendation

• There are also another types of relations which can improve the recommendation performance.
  • Other Relations between users and items.
A Real Example

- **Recommendation Task:** Recommend Conference for Author.
- **Target Recommendation Relation:** 
  \( \text{Author-Paper-Conference} \)
- **Social Information for improvement:**
  - Relations between users: Author-Paper-Author.
  - Relations between items: Conference-Topic-Conference.
  - Relations between users and items: Author-Topic-Author-Paper-Conf.
- We want to design a model to use all these three types of relation.
Outline

• Background and Motivation

• The Hete-CF Model and Algorithm

• Experimental Result

• Conclusion
Model Overview

• Part 1: Modeling Relations between users
  • A good recommendation result should consider the relations between users.

• Part 2: Modeling Relations between items
  • A good recommendation result should consider the relations between items.

• Part 3: Modeling Relations between users and items
  • Some other relations between users and items can improve the recommendation result.

A Unified Model.
Part 1: Modeling Relations between users

- Using Graph Regularization:

\[
\min_{U,A} \sum_{k=0}^{N_A} \alpha_k \sum_{i=0}^{n} \sum_{j=0}^{n} S^k_A(i, j) \|U_i - U_j\|_F^2
\]

- **A**: weight for each relations between users introduced.
- **\( S^k_A \)** is the similarity matrix between users, can be calculated as follow:
  \[
  S^k_A(i, j) = S^{PS}_{PA}(i, j)
  \]
- The **\( PA_{k}^A \)** can be calculated using **PathSim (VLDB’11)**
Part 2: Modeling Relations between items

• Using Graph Regularization:

\[
\min_{V_i, B} \sum_{k=0}^{N_B} \beta_k \sum_{i=0}^{m} \sum_{j=0}^{m} S_B^k(i, j) \|V_i - V_j\|_F^2
\]

• **B**: weight for each relation between items introduced.

• \(S_B^k\) is the similarity matrix between items, can be calculated as follow:

\[
S_B^k(i, j) = S_{P_{k}}^{PS}(i, j)
\]

• The \(S_{P_{k}}^{PS}(i, j)\) can be calculated using **PathSim**
Part 3: Modeling Relations between users and items

• Using Collaborative Model:

\[
\min_{U,V,W} \sum_{k=0}^{N_W} w_k \sum_{i=0}^{n} \sum_{j=0}^{m} (U^T_i V_j - R^k_{i,j})^2
\]

• \( W \): weight for each relations between users and items introduced.

• \( R^k \) is the relation graph of the \( k-th \) relation between users and items:

\[
R^k(i,j) = S_{P^W_k}^{PS}(i,j)
\]

• The \( S_{P^W_k}^{PS}(i,j) \) can be calculated using PathSim
The Unified Overview

• Model

\[
\begin{align*}
\min_{U,V,A,B,W} & \sum_{i=0}^{m} \sum_{j=0}^{n} (U^T_i V_j - R_{i,j})^2 \\
+ & \sum_{k=0}^{N_A} \alpha_k \sum_{i=0}^{n} \sum_{j=0}^{n} S_A^k(i,j) \|U_i - U_j\|_F^2 \\
+ & \sum_{k=0}^{N_B} \beta_k \sum_{i=0}^{m} \sum_{j=0}^{m} S_B^k(i,j) \|V_i - V_j\|_F^2 \\
+ & \sum_{k=0}^{N_W} \gamma_k \sum_{i=0}^{n} \sum_{j=0}^{m} (U^T_i V_j - R_{i,j}^k)^2 \\
+ & \lambda(\|U\|_F^2 + \|V\|_F^2 + \|A\|_F^2 + \|B\|_F^2 + \|W\|_F^2)
\end{align*}
\]

- Modeling the relations between users
- Modeling the relations between items
- Modeling the relations between users and items
The Learning Algorithm

- A Two-Step Iteration Method.
  - The predicted rating vector $U, V$ and the weight for each relation $A, B, W$ mutually enhance each other.

- Optimize $U, V$ Given $A, B, W$
  - Becomes to a traditional CF model
  - Using SGD to obtain $U, V$

- Optimize $A, B, W$ Given $U, V$
  - Becomes a linear model
  - Also using SGD to obtain $A, B, W$

---

Input: A heterogeneous information network $G = (O, E, W)$.
Three sets of meta-paths between user and item, users, and items. The user-item rating matrix $R$. Parameter $\lambda$, $\alpha$.

Output: The rating Matrix $\hat{R}$;

1. Initialize $U, V, A, B, W$ randomly;
2. while not reaching the inner $U, V, A, B, C$ difference threshold do
   3. while not reaching the inner $U, V$ difference threshold do
      4. Update $U, V$ using Eqs. (16) and (17);
   5. end
   6. while not reaching the inner $A, B, C$ difference threshold do
      7. Update $A, B, W$ using Eqs. (22), (23) and (24);
   8. end
3. end
4. The prediction rating is $\hat{R}_{ij} = f(U^T_i V_j)$;
5. return $\hat{R}_{ij}$;

Algorithm 1: Hete-CF
Outline

• Background and Motivation

• The Hete-CF Model and Algorithm

• Experimental Result

• Conclusion
Experimental Setup

- **Datasets**
  - Sub-DBLP Dataset
  - 261 Journal and 313 conference paper

- **Task**
  - Recommend *Conference* to *Author*

- **Cross Validation**
  - 5-fold cross validation
  - 40%(60%) as training
  - 60%(40%) as testing

- **Baseline**
  - UserMean, ItemMean, NMF

- **State-of-the-art**
  - Trust-MF (IJCAI-13),
  - Hete-MF (IJCAI-HINA 13)

- **Evaluation**
  - MAE
  - RMSE
Experimental Result

- Result

<table>
<thead>
<tr>
<th>% Training</th>
<th>Feature</th>
<th>Evaluation</th>
<th>UserMean</th>
<th>ItemMean</th>
<th>NMF</th>
<th>Trust-MF</th>
<th>Hete-MF</th>
<th>Hete-CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>40%</td>
<td>d = 5</td>
<td>MAE</td>
<td>0.942 ± 0.02</td>
<td>1.065 ± 0.02</td>
<td>2.156 ± 0.02</td>
<td>0.831 ± 0.01</td>
<td>0.931 ± 0.02</td>
<td>0.831 ± 0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>1.216 ± 0.01</td>
<td>1.123 ± 0.02</td>
<td>2.394 ± 0.01</td>
<td>1.013 ± 0.02</td>
<td>1.105 ± 0.01</td>
<td>1.002 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>d = 10</td>
<td>MAE</td>
<td>0.943 ± 0.03</td>
<td>0.948 ± 0.01</td>
<td>2.194 ± 0.03</td>
<td>0.887 ± 0.01</td>
<td>0.901 ± 0.01</td>
<td>0.859 ± 0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>1.138 ± 0.02</td>
<td>1.256 ± 0.04</td>
<td>2.292 ± 0.02</td>
<td>1.083 ± 0.03</td>
<td>1.114 ± 0.03</td>
<td>1.056 ± 0.02</td>
</tr>
<tr>
<td>60%</td>
<td>d = 5</td>
<td>MAE</td>
<td>0.948 ± 0.02</td>
<td>0.919 ± 0.01</td>
<td>2.131 ± 0.04</td>
<td>0.812 ± 0.02</td>
<td>0.891 ± 0.02</td>
<td>0.831 ± 0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>1.132 ± 0.02</td>
<td>1.157 ± 0.01</td>
<td>2.385 ± 0.01</td>
<td>0.907 ± 0.02</td>
<td>1.010 ± 0.03</td>
<td>0.938 ± 0.02</td>
</tr>
<tr>
<td></td>
<td>d = 10</td>
<td>MAE</td>
<td>0.932 ± 0.03</td>
<td>0.978 ± 0.03</td>
<td>2.184 ± 0.02</td>
<td>0.873 ± 0.03</td>
<td>0.881 ± 0.01</td>
<td>0.856 ± 0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>1.154 ± 0.02</td>
<td>1.143 ± 0.02</td>
<td>2.275 ± 0.01</td>
<td>1.051 ± 0.01</td>
<td>1.013 ± 0.02</td>
<td>0.994 ± 0.03</td>
</tr>
</tbody>
</table>
Outline

• Background and Motivation

• The Hete-CF Model and Algorithm

• Experimental Result

• Conclusion
Conclusion

• We proposed Hete-CF a CF-based social recommendation method using heterogeneous relations

• Hete-CF utilize three types of relations in a heterogeneous network.

• The experimental result showed the effectiveness of Hete-CF
Chen Luo
Master Candidate from Jilin University
Home Page: http://www.chen-luo.com/