Recommender Systems with Social Regularization

Hao Ma, Dengyong Zhou, Chao Liu
Microsoft Research

Michael R. Lyu
The Chinese University of Hong Kong

Irwin King
AT&T Labs – Research
The Chinese University of Hong Kong
Recommender Systems are Everywhere
Web 2.0 Web Sites are Everywhere
Trust-aware Recommender Systems

• These Methods utilize the inferred implicit or observed explicit trust information to further improve traditional recommender systems.
  – [P. Massa and P. Avesani, RecSys 2007]
  – [H. Ma, I. King, and M. R. Lyu, SIGIR 2009]

• Based on the motivation that “I trust you => I have similar tastes with you”.
Comparison

• Trust-aware

  – Trust network: unilateral relations

  – Trust relations can be treated as “similar” relations

  – Few datasets available on the Web

• Social-based

  – Social friend network: mutual relations

  – Friends are very diverse, and may have different tastes

  – Lots of Web sites have social network implementation
Contents of This Work

• Focusing on social-based recommendation problems

• Two methods are proposed based on matrix factorization with social regularization terms
  – Can be applied to trust-aware recommender systems.

• Experiments on two large datasets
  – Douban (social friend network)
  – Epinions (trust network)
Problem Definition

Social Network Information

User-Item Rating Matrix
Low-Rank Matrix Factorization for Collaborative Filtering

• Objective function

$$\min_{U,V} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij} (R_{ij} - U_i^T V_j)^2 + \frac{\lambda_1}{2} ||U||_F^2 + \frac{\lambda_2}{2} ||V||_F^2$$

$I_{ij}$ is the indicator function that is equal to 1 if user $u_i$ rated item $v_j$ and equal to 0 otherwise.

$U, V$: low dimension column vectors to represent user/item preferences.
Social Regularization I

- Average-based regularization

$$\min_{U,V} \mathcal{L}_1(R, U, V) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{i,j} (R_{i,j} - U_i^T V_j)^2$$

$$+ \frac{\alpha}{2} \sum_{i=1}^{m} \|U_i - \frac{\sum_{f \in \mathcal{F}+(i)} Sim(i, f) \times U_f}{\sum_{f \in \mathcal{F}+(i)} Sim(i, f)} \|^2_F$$

$$+ \frac{\lambda_1}{2} \|U\|^2_F + \frac{\lambda_2}{2} \|V\|^2_F.$$

Minimize $U_i$'s taste with the average tastes of $U_i$'s friends. The similarity function $Sim(i, f)$ allows the social regularization term to treat users’ friends differently.
Social Regularization I

- Gradients

\[
\frac{\partial L_1}{\partial U_i} = \sum_{j=1}^{n} I_{ij} (U_i^T V_j - R_{ij}) V_j + \lambda_1 U_i \\
+ \alpha (U_i - \frac{\sum_{f \in \mathcal{F}^+(i)} Sim(i, f) \times U_f}{\sum_{f \in \mathcal{F}^+(i)} Sim(i, f)}) \\
+ \alpha \sum_{g \in \mathcal{F}^+(i)} \frac{-Sim(i, g) (U_g - \frac{\sum_{f \in \mathcal{F}^+(g)} Sim(g, f) \times U_f}{\sum_{f \in \mathcal{F}^+(g)} Sim(g, f)})}{\sum_{f \in \mathcal{F}^+(g)} Sim(g, f)}
\]

\[
\frac{\partial L_1}{\partial V_j} = \sum_{i=1}^{m} I_{ij} (U_i^T V_j - R_{ij}) U_i + \lambda_2 V_j.
\]
Social Regularization II

• Individual-based regularization

\[
\min_{U,V} \mathcal{L}_2(R, U, V) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij} (R_{ij} - U_i^T V_j)^2 \\
+ \frac{\beta}{2} \sum_{i=1}^{m} \sum_{f \in \mathcal{F}^+(i)} \text{Sim}(i, f) \|U_i - U_f\|_F^2 \\
+ \lambda_1 \|U\|_F^2 + \lambda_2 \|V\|_F^2.
\]

This approach allows similarity of friends’ tastes to be individually considered. It also indirectly models the propagation of tastes.
Social Regularization II

• Gradients

\[
\frac{\partial L_2}{\partial U_i} = \sum_{j=1}^{n} I_{ij} (U_i^T V_j - R_{ij}) V_j + \lambda_1 U_i \\
+ \beta \sum_{f \in F^+(i)} Sim(i, f) (U_i - U_f) \\
+ \beta \sum_{g \in F^-(i)} Sim(i, g) (U_i - U_g),
\]

\[
\frac{\partial L_2}{\partial V_j} = \sum_{i=1}^{m} I_{ij} (U_i^T V_j - R_{ij}) U_i + \lambda_2 V_j.
\]
Similarity Function

• Vector Space Similarity (VSS) or Cosine Similarity

\[
Sim(i, f) = \frac{\sum_{j \in I(i) \cap I(f)} R_{ij} \cdot R_{fj}}{\sqrt{\sum_{j \in I(i) \cap I(f)} R_{ij}^2} \cdot \sqrt{\sum_{j \in I(i) \cap I(f)} R_{fj}^2}}
\]

• Pearson Correlation Coefficient (PCC)

\[
Sim(i, f) = \frac{\sum_{j \in I(i) \cap I(f)} (R_{ij} - \bar{R}_i) \cdot (R_{fj} - \bar{R}_f)}{\sqrt{\sum_{j \in I(i) \cap I(f)} (R_{ij} - \bar{R}_i)^2} \cdot \sqrt{\sum_{j \in I(i) \cap I(f)} (R_{fj} - \bar{R}_f)^2}}
\]
Dataset I

• Douban
  – Chinese Web 2.0 Web site with social friend network service
  – The largest online book, movie and music review and rating site in China
  – We crawled 129,940 users and 58,541 movies with 16,830,839 movie ratings
  – The total number of friend links between users is 1,692,952
Dataset II

• Epinions
  – A well-known English general consumer review and rating site
  – Every member maintains a trust list which presents a user network of trust relationships
  – We crawled 51,670 users who have rated a total of 83,509 different items. The total number of ratings is 631,064
  – The total number of issued trust statements is 511,799
Metrics

- MAE and RMSE

\[
MAE = \frac{1}{T} \sum_{i,j} |R_{ij} - \hat{R}_{ij}|
\]

\[
RMSE = \sqrt{\frac{1}{T} \sum_{i,j} (R_{ij} - \hat{R}_{ij})^2}
\]
# Performance Comparison

## Table 5: Performance Comparisons (Dimensionality = 10)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Training</th>
<th>Metrics</th>
<th>UserMean</th>
<th>ItemMean</th>
<th>NMF</th>
<th>PMF</th>
<th>RSTE</th>
<th>SRT\textsubscript{vss}</th>
<th>SRT\textsubscript{pcc}</th>
<th>SR2\textsubscript{vss}</th>
<th>SR2\textsubscript{pcc}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Douban</strong></td>
<td>80%</td>
<td>MAE</td>
<td>0.6809</td>
<td>0.6288</td>
<td>0.5732</td>
<td>0.5693</td>
<td>0.5643</td>
<td>0.5579</td>
<td>0.5576</td>
<td>0.5548</td>
<td>0.5543</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>18.59%</td>
<td>11.85%</td>
<td>3.30%</td>
<td>2.63%</td>
<td>1.77%</td>
<td>0.7026</td>
<td>0.7022</td>
<td>0.6992</td>
<td>0.6988</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>0.8480</td>
<td>0.7898</td>
<td>0.7225</td>
<td>0.7200</td>
<td>0.7144</td>
<td>0.7081</td>
<td>0.7078</td>
<td>0.7046</td>
<td>0.7042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>17.59%</td>
<td>11.52%</td>
<td>3.28%</td>
<td>2.94%</td>
<td>2.18%</td>
<td>0.7081</td>
<td>0.7078</td>
<td>0.7046</td>
<td>0.7042</td>
</tr>
<tr>
<td></td>
<td>60%</td>
<td>MAE</td>
<td>0.6823</td>
<td>0.6300</td>
<td>0.5768</td>
<td>0.5737</td>
<td>0.5698</td>
<td>0.5627</td>
<td>0.5623</td>
<td>0.5597</td>
<td>0.5593</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>18.02%</td>
<td>11.22%</td>
<td>3.03%</td>
<td>2.51%</td>
<td>1.84%</td>
<td>0.7081</td>
<td>0.7078</td>
<td>0.7046</td>
<td>0.7042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>0.8505</td>
<td>0.7926</td>
<td>0.7351</td>
<td>0.7290</td>
<td>0.7207</td>
<td>0.7081</td>
<td>0.7078</td>
<td>0.7046</td>
<td>0.7042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>17.20%</td>
<td>11.15%</td>
<td>4.20%</td>
<td>3.40%</td>
<td>2.29%</td>
<td>0.7081</td>
<td>0.7078</td>
<td>0.7046</td>
<td>0.7042</td>
</tr>
<tr>
<td></td>
<td>40%</td>
<td>MAE</td>
<td>0.6854</td>
<td>0.6317</td>
<td>0.5899</td>
<td>0.5868</td>
<td>0.5767</td>
<td>0.5706</td>
<td>0.5702</td>
<td>0.5690</td>
<td>0.5685</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>17.06%</td>
<td>10.00%</td>
<td>3.63%</td>
<td>3.12%</td>
<td>1.42%</td>
<td>0.7172</td>
<td>0.7169</td>
<td>0.7129</td>
<td>0.7125</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>0.8567</td>
<td>0.7971</td>
<td>0.7482</td>
<td>0.7411</td>
<td>0.7295</td>
<td>0.7172</td>
<td>0.7169</td>
<td>0.7129</td>
<td>0.7125</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>16.83%</td>
<td>10.61%</td>
<td>4.77%</td>
<td>3.86%</td>
<td>2.33%</td>
<td>0.7172</td>
<td>0.7169</td>
<td>0.7129</td>
<td>0.7125</td>
</tr>
<tr>
<td><strong>Epinions</strong></td>
<td>90%</td>
<td>MAE</td>
<td>0.9134</td>
<td>0.9768</td>
<td>0.8712</td>
<td>0.8651</td>
<td>0.8367</td>
<td>0.8290</td>
<td>0.8287</td>
<td>0.8258</td>
<td>0.8256</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>9.61%</td>
<td>15.48%</td>
<td>5.23%</td>
<td>4.57%</td>
<td>1.33%</td>
<td>1.0792</td>
<td>1.0790</td>
<td>1.0744</td>
<td>1.0739</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>1.1688</td>
<td>1.2375</td>
<td>1.1621</td>
<td>1.1544</td>
<td>1.1094</td>
<td>1.0792</td>
<td>1.0790</td>
<td>1.0744</td>
<td>1.0739</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>8.12%</td>
<td>13.22%</td>
<td>7.59%</td>
<td>6.97%</td>
<td>3.20%</td>
<td>1.0792</td>
<td>1.0790</td>
<td>1.0744</td>
<td>1.0739</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>MAE</td>
<td>0.9285</td>
<td>0.9913</td>
<td>0.8951</td>
<td>0.8886</td>
<td>0.8537</td>
<td>0.8493</td>
<td>0.8491</td>
<td>0.8447</td>
<td>0.8443</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>9.07%</td>
<td>14.83%</td>
<td>5.68%</td>
<td>4.99%</td>
<td>1.10%</td>
<td>1.1016</td>
<td>1.1013</td>
<td>1.0958</td>
<td>1.0954</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>1.1817</td>
<td>1.2584</td>
<td>1.1832</td>
<td>1.1760</td>
<td>1.1256</td>
<td>1.1016</td>
<td>1.1013</td>
<td>1.0958</td>
<td>1.0954</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve</td>
<td>7.30%</td>
<td>12.95%</td>
<td>7.42%</td>
<td>6.85%</td>
<td>2.68%</td>
<td>1.1016</td>
<td>1.1013</td>
<td>1.0958</td>
<td>1.0954</td>
</tr>
</tbody>
</table>
Impact of Parameter $\beta$

(a) Douban (MAE)

(b) Douban (RMSE)

(c) Epinions (MAE)

(d) Epinions (RMSE)
# Impact of Similarity Functions

## Table 6: Similarity Analysis (Dimensionality = 10)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Training</th>
<th>Metrics</th>
<th>SR2 Sim=1</th>
<th>SR2 Sim=Ran</th>
<th>SR2_vss</th>
<th>SR2_pcc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Douban</td>
<td>80%</td>
<td>MAE</td>
<td>0.5579</td>
<td>0.5592</td>
<td>0.5548</td>
<td>0.5543</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>0.7034</td>
<td>0.7047</td>
<td>0.6992</td>
<td>0.6988</td>
</tr>
<tr>
<td></td>
<td>60%</td>
<td>MAE</td>
<td>0.5631</td>
<td>0.5643</td>
<td>0.5597</td>
<td>0.5593</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>0.7083</td>
<td>0.7098</td>
<td>0.7046</td>
<td>0.7042</td>
</tr>
<tr>
<td></td>
<td>40%</td>
<td>MAE</td>
<td>0.5724</td>
<td>0.5737</td>
<td>0.5690</td>
<td>0.5685</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>0.7195</td>
<td>0.7209</td>
<td>0.7129</td>
<td>0.7125</td>
</tr>
<tr>
<td>Epinions</td>
<td>90%</td>
<td>MAE</td>
<td>0.8324</td>
<td>0.8345</td>
<td>0.8258</td>
<td>0.8256</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>1.0794</td>
<td>1.0809</td>
<td>1.0744</td>
<td>1.0739</td>
</tr>
<tr>
<td></td>
<td>80%</td>
<td>MAE</td>
<td>0.8511</td>
<td>0.8530</td>
<td>0.8447</td>
<td>0.8443</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>1.1002</td>
<td>1.1018</td>
<td>1.0958</td>
<td>1.0954</td>
</tr>
</tbody>
</table>
Conclusions

- We proposed two new social recommendation methods
- Our approaches perform better than other traditional and trust-aware recommendation methods
- The methods scale well since the employed algorithm is linear with the observation of ratings
Future Work

• Employ more accurate similarity functions

• Consider item side regularization

• Apply similar techniques to other social applications, like social search problems
Thanks!

Q&A