New Objective Functions for Social Collaborative Filtering

Joseph Noel, Scott Sanner, Khoi-Nguyen Tran, Peter Christen, Lexing Xie, Edwin Bonilla, Ehsan Abbasnejad, Nicolas Della Penna

NICTA

THE AUSTRALIAN NATIONAL UNIVERSITY
ANU Link Recommender (LinkR)

- Recommends 3 daily links on Facebook

  - Non-friend Recommendation (only link context)
  - Rating + Optional Link Feedback
  - Friend Recommendation (friend message + link context)
ANU Link Recommender (LinkR)

- Recommends 3 daily links on Facebook

This talk:
- Existing baselines
- New algorithms (training objectives)
- Live user trial evaluation (5 months)
- Lessons learned and future work

How to leverage social network and App data to learn to recommend links?

Non-friend Recommendation (only link context)

Friend Recommendation (friend message + link context)
Recommendation

• Predict **missing** from **observed** ratings?

\[
R = \begin{pmatrix}
1 & 1 & 1 & 0 & ? & 0 \\
1 & 0 & ? & 0 & 0 & 0 \\
0 & 0 & 1 & ? & & \\
\end{pmatrix}
\]

**Canonical Example:**
Netflix Competition

...1-5 ratings, here: like (1), dislike (0)
Social Recommendation

- Adds indirect social context to users

Main question in this work:
How to incorporate social context to improve predictions?
Question 1:

What existing (social) collaborative filtering techniques make good Facebook link recommenders?
Content-based Filtering (CBF)

- Predict like / dislike directly from features

\[
R = \begin{pmatrix}
1 & 1 & 1 & 0 & ? & 0 \\
1 & 0 & ? & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 1 & ? & \end{pmatrix}
\]

\[R(\text{user } x, \text{ movie } y) = f(\Phi_x, \Phi_y)\]

Trained classifier, e.g. SVM

Sci-Fi, Director: Mel Brooks

Romance, Starring: Julia Roberts, Richard Gere

29, Male, Sydney

24, Male, Canberra

33, Male, Canberra
Collaborative Filtering (CF): KNN

- No features? k-nearest neighbor, e.g., \( k = 2 \)

\[
R = \begin{bmatrix}
1 & 1 & 1 & 0 & \? & 0 \\
1 & 0 & \? & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 1 & \? & \?
\end{bmatrix}
\]

\[
\frac{.71 \times 1 + .50 \times 0}{.71 + .50} = 0.59
\]

\[
0.71
\]

\[
0.50
\]
Collaborative Filtering: PMF

• Or low $k$-rank matrix factorization, e.g. $k=2$

\[ R = \begin{pmatrix}
1 & 1 & 1 & 0 & ? & 0 \\
1 & 0 & ? & 0 & 0 & \\
0 & 0 & 1 & ? & \\
\end{pmatrix} = \begin{pmatrix}
.7 & .3 & \\
.5 & -.2 & \\
-.1 & .9 & \\
\end{pmatrix} \begin{pmatrix}
\cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot \\
\end{pmatrix}
\]

Latent User Features

Latent Item Features

\[ \min_{U, V} \sum_{(x, y) \in D} \frac{1}{2} (R_{x,y} - U^T x V_y)^2 \]

Standard PMF CF Objective (reg. not shown), novel objectives build on this.
Features in CF: **Matchbox**

\[
R = \begin{pmatrix}
1 & 1 & 1 & 0 & ? & 0 \\
1 & 0 & ? & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 1 & ? & & \\
\end{pmatrix}
\]

\[
(Ux)^T \begin{pmatrix}
.7 & .3 \\
.5 & -.2 \\
.3 & -.7 \\
\end{pmatrix} \begin{pmatrix}
V_y \\
\end{pmatrix}
\]

Project features into latent space – helps cold-start problem.

Reduces to previous PMF CF if \(x, y\) are indicators.
Social Collaborative Filtering

\[ R = \begin{pmatrix}
1 & 1 & 1 & 0 & ? & 0 \\
1 & 0 & ? & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 1 & ? & \vdots & \vdots \\
\end{pmatrix} \]

\[ \text{Int}_{x,z} = \frac{1}{N(N-1)} \sum_{x',z' \neq x} \text{# interactions by } x \\
S_{x,z} = \ln (\text{Int}_{x,z}) \]

\[ \text{PMF + Social Regularization} \]

\[ \min_U \sum_x \sum_{z \in \text{friends}_x} \frac{1}{2} (S_{x,z} - \langle U_x, U_z \rangle)^2 \]

\[ \text{PMF + Social Spectral Reg.} \]

\[ \min_U \sum_x \sum_{z \in \text{friends}_x} \frac{1}{2} S_{x,z}^+ \| U_x - U_z \|_2^2 \]

A totally awesome link!
Like!
Like!
Extra cute baby!
Check out this awesome link!
Like!
Like!
Extra cute baby!
Another awesome link!
Like!
A totally awesome link!
We find that social CF based on matrix factorization works best.

Question 2:

Can we extend objectives to make it work better?

Yes, we introduce new objective functions for social collaborative filtering.
Objective Framework

\[
\min_{w, U, V} \text{Obj} = \sum_{i} \lambda_i \text{Obj}_i
\]

Standard Objective

\[
\text{Obj}_{pmcf} = \sum_{(x,y) \in D} \frac{1}{2} (R_{x,y} - [\sigma] x^T U^T V y)^2
\]

Standard Regularizers

\[
\text{Obj}_{ru} = \frac{1}{2} \|U\|_{\text{Fro}}^2 = \frac{1}{2} \text{tr}(U^T U)
\]

\[
\text{Obj}_{rv} = \frac{1}{2} \text{tr}(V^T V)
\]

\[
\text{Obj}_{rw} = \frac{1}{2} \|w\|_2^2 = \frac{1}{2} w^T w
\]

Social Regularizers

\[
\text{Obj}_{rs} = \sum_{x} \sum_{z \in \text{friends}(x)} \frac{1}{2} (S_{x,z} - \langle U x, U z \rangle)^2
\]

Prediction objectives and regularizers to constrain learning.

Other predictors aside from MF?

This is first proposal... feature-based S.R.

Other social regularizers?
Proposal 1 ½

- Use interactions to learn latent spectral projection of **user and features**

\[
\text{Obj}_{rss} = \sum_{x} \sum_{z \in \text{friends}(x)} \frac{1}{2} S_{x,z}^{+} \| Ux - Uz \|_{2}^{2}
\]

\[
= \sum_{x} \sum_{z \in \text{friends}(x)} \frac{1}{2} S_{x,z}^{+} (x - z)^{T} U^{T} U (x - z)
\]

Don’t predict $S_{x,z}$, use it to vary regularization strength!
Proposal II

– Directly model information diffusion

\[
Obj_{phy} = \sum_{(x,y) \in D} \frac{1}{2} (R_{x,y} - [\sigma]w^{T}f_{x,y} - [\sigma]x^{T}U^{T}V y)^{2}
\]

Features such as:
Did user z (a friend of x), also like y?
Proposal III

- Exploit the fact that users have common interests in restricted areas
  - Use co-preferences $P_{x,z,y}$
    - Did users $x$ and $z$ (dis)like item $y$?

$$Obj_{cp} = \sum_{(x,z,y) \in C} \frac{1}{2} \left( P_{x,z,y} - \langle Ux, Uz \rangle_{Vy} \right)^2$$

$$= \sum_{(x,z,y) \in C} \frac{1}{2} \left( P_{x,z,y} - x^T U^T \text{diag}(Vy) U z \right)^2$$

- And also spectral variant

Reweight user regularization according to latent dimensions for co-preferred item.
Social Recommendation Evaluation via User Trials

Link Recommendation on Facebook
ANU Link Recommender (LinkR)

- Recap: Recommend 3 daily links on Facebook

Non-friend Recommendation (only link context)

Rating + Optional Link Feedback

Friend Recommendation (friend message + link context)
Trials and Algorithms

• **Trial 1: Baselines**
  – **SVM** (Content-based filtering – CBF)
  – **KNN** (Collaborative filtering – CF)
  – Matchbox – **MB** (CF + CBF)
  – Social Matchbox – **SMB** (CBF + CF + Soc. Reg)

• **Trial 2: New Objectives**
  – **SMB**
  – Spectral Reg. variant of SMB – **Sp. MB**
  – SMB + Information Diffusion – **S. Hybrid**
  – MB + Spectral Copreference Reg. – **S. CP**
# LinkR Statistics

## Table

<table>
<thead>
<tr>
<th>Column</th>
<th>#Records (App Users)</th>
<th>#Records (App User and Friends)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users</td>
<td>103</td>
<td>39,850</td>
</tr>
<tr>
<td>Gender</td>
<td>102</td>
<td>36,401</td>
</tr>
<tr>
<td>Birthday</td>
<td>103</td>
<td>27,624</td>
</tr>
</tbody>
</table>

## Breakdown

<table>
<thead>
<tr>
<th>Breakdown</th>
<th>Count (App Users)</th>
<th>Count (App User and Friends)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>73</td>
<td>19,742</td>
</tr>
<tr>
<td>Female</td>
<td>29</td>
<td>16,659</td>
</tr>
<tr>
<td>High School</td>
<td>104</td>
<td>29,503</td>
</tr>
<tr>
<td>College</td>
<td>115</td>
<td>29,223</td>
</tr>
<tr>
<td>Graduate School</td>
<td>56</td>
<td>7733</td>
</tr>
</tbody>
</table>

## App Users

<table>
<thead>
<tr>
<th>App Users</th>
<th>Posts</th>
<th>Tags</th>
<th>Comments</th>
<th>Likes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wall</td>
<td>27,955</td>
<td>5,256</td>
<td>15,121</td>
<td>11,033</td>
</tr>
<tr>
<td>Link</td>
<td>3,974</td>
<td>—</td>
<td>5,757</td>
<td>4,279</td>
</tr>
<tr>
<td>Photo</td>
<td>4,147</td>
<td>22,633</td>
<td>8,677</td>
<td>5,938</td>
</tr>
<tr>
<td>Video</td>
<td>211</td>
<td>2,105</td>
<td>1,687</td>
<td>710</td>
</tr>
</tbody>
</table>

## App Users and Friends

<table>
<thead>
<tr>
<th>App Users and Friends</th>
<th>Posts</th>
<th>Tags</th>
<th>Comments</th>
<th>Likes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wall</td>
<td>3,384,740</td>
<td>912,687</td>
<td>2,152,321</td>
<td>1,555,225</td>
</tr>
<tr>
<td>Link</td>
<td>514,475</td>
<td>—</td>
<td>693,930</td>
<td>666,631</td>
</tr>
<tr>
<td>Photo</td>
<td>1,098,679</td>
<td>8,407,822</td>
<td>2,978,635</td>
<td>1,960,138</td>
</tr>
<tr>
<td>Video</td>
<td>56,241</td>
<td>858,054</td>
<td>463,401</td>
<td>308,763</td>
</tr>
</tbody>
</table>
LinkR Usage Statistics

### Trial 1 – Aug. 25, 2011 to Oct. 13, 2011

<table>
<thead>
<tr>
<th></th>
<th>SMB</th>
<th>MB</th>
<th>SVM</th>
<th>KNN</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users All</td>
<td>26</td>
<td>26</td>
<td>28</td>
<td>28</td>
<td>108</td>
</tr>
<tr>
<td>Users ≥ 10</td>
<td>13</td>
<td>9</td>
<td>13</td>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>Users ≥ 30</td>
<td>9</td>
<td>3</td>
<td>11</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>Ratings All</td>
<td>819</td>
<td>526</td>
<td>901</td>
<td>242</td>
<td>2508</td>
</tr>
<tr>
<td>Ratings ≥ 10</td>
<td>811</td>
<td>505</td>
<td>896</td>
<td>228</td>
<td>2440</td>
</tr>
<tr>
<td>Ratings ≥ 30</td>
<td>737</td>
<td>389</td>
<td>851</td>
<td>182</td>
<td>2159</td>
</tr>
<tr>
<td>Clicks All</td>
<td>383</td>
<td>245</td>
<td>413</td>
<td>218</td>
<td>1259</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th></th>
<th>SMB</th>
<th>Sp.MB</th>
<th>Sp.CP</th>
<th>SHyb.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users All</td>
<td>27</td>
<td>27</td>
<td>29</td>
<td>28</td>
<td>111</td>
</tr>
<tr>
<td>Users ≥ 10</td>
<td>15</td>
<td>11</td>
<td>8</td>
<td>12</td>
<td>46</td>
</tr>
<tr>
<td>Users ≥ 30</td>
<td>12</td>
<td>9</td>
<td>5</td>
<td>10</td>
<td>36</td>
</tr>
<tr>
<td>Ratings All</td>
<td>1434</td>
<td>882</td>
<td>879</td>
<td>614</td>
<td>3809</td>
</tr>
<tr>
<td>Ratings ≥ 10</td>
<td>1411</td>
<td>878</td>
<td>863</td>
<td>602</td>
<td>3754</td>
</tr>
<tr>
<td>Ratings ≥ 30</td>
<td>1348</td>
<td>850</td>
<td>802</td>
<td>570</td>
<td>3570</td>
</tr>
<tr>
<td>Clicks All</td>
<td>553</td>
<td>320</td>
<td>278</td>
<td>199</td>
<td>1350</td>
</tr>
</tbody>
</table>
Trial 1: Baselines

Likes (dark) over Dislikes (light)

Friends vs. Non-friends

Ratio of Liked Ratings to Disliked Ratings Per Algorithm

Ratio of Liked to Disliked Recommendations From Friends vs. Non-Friends
Trial 2: New Objectives

Friends

Non-friends
Click Behavior

Ratings for Clicked Links

Clicks vs. Description

Ratings vs. Description
Impact of Popularity
## Individual Link Comments

<table>
<thead>
<tr>
<th>Comment Type</th>
<th>#</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>not interested</td>
<td>88</td>
<td>36.5%</td>
</tr>
<tr>
<td>wrong language</td>
<td>37</td>
<td>15.4%</td>
</tr>
<tr>
<td>really liked it!</td>
<td>35</td>
<td>14.5%</td>
</tr>
<tr>
<td>bad YouTube</td>
<td>25</td>
<td>10.4%</td>
</tr>
<tr>
<td>seen it already</td>
<td>25</td>
<td>10.4%</td>
</tr>
<tr>
<td>problem / dead</td>
<td>20</td>
<td>8.3%</td>
</tr>
<tr>
<td>outdated</td>
<td>7</td>
<td>2.9%</td>
</tr>
<tr>
<td>miscellaneous</td>
<td>4</td>
<td>1.7%</td>
</tr>
</tbody>
</table>
# Survey

<table>
<thead>
<tr>
<th>User Survey Comments</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>want more control over recommendations made (music, blogs, news)</td>
<td></td>
</tr>
<tr>
<td>want option to see &gt; 3 recommendations</td>
<td></td>
</tr>
<tr>
<td>links need description / context or explanation of recommendation</td>
<td></td>
</tr>
<tr>
<td>more variety, diversity</td>
<td></td>
</tr>
</tbody>
</table>
Experimental Design in Retrospect

• Originally wanted to do **active learning**
  – In our Google Grant proposal
  – But with user uptake, difficult to evaluate this
    • Need very active users (only 25% were active)

• But we were stuck with the original experimental design after the first trial
  • Hard to statistically compare small user groups

• If do again, would instead interleave interactions (each recommendation comes from a randomly selected algorithm – so *all* users see *all* algorithms)

• But main results of Spec. MB significant nonetheless
Conclusions

• **Social spectral regularization**
  – Undeniably the top-performer
  – As good as direct information diffusion features
  – Interactions stronger than co-preferences
    • Or co-preferences harder to optimize?

• **Overall**
  – Machine learning works!
    • Better than more ad-hoc methods like KNN
    • Power of latent factorization methods
  – User socially informed regularizers!
    • In general, users who interact a lot have similar preferences!
Future Work

• Are all interactions equal?

• No!
  – Learning predictiveness of fine-grained interactions can do as well as MF, but with simple classifiers!
  – Work in progress…
Thank you!

Especially to Doug Aberdeen (Google Zurich) for supporting our Google Grant

And to Sally-Ann Williams for 100+ pairs of Google flip-flops, which have helped us attract users for our live trials!
Extra Slides
Aside: Matrix Definitions

\[
U = \begin{bmatrix}
U_{1,1} & \cdots & U_{1,I} \\
\vdots & \ddots & \vdots \\
U_{K,1} & \cdots & U_{K,I}
\end{bmatrix}
\]

\[
V = \begin{bmatrix}
V_{1,1} & \cdots & V_{1,J} \\
\vdots & \ddots & \vdots \\
V_{K,1} & \cdots & V_{K,J}
\end{bmatrix}
\]
Proposal I

• Use interactions to learn latent projection of user \textbf{and features}

\[
\text{Obj}_{rs} = \sum_{x} \sum_{z \in \text{friends}(x)} \frac{1}{2} (S_{x,z} - \langle Ux, Uz \rangle)^2 \\
= \sum_{x} \sum_{z \in \text{friends}(x)} \frac{1}{2} (S_{x,z} - x^T U^T U z)^2
\]