Social Recommendation

Irwin King with Hao Ma

ATT Labs Research
&
Department of Computer Science and Engineering
The Chinese University of Hong Kong

king@cse.cuhk.edu.hk
http://www.cse.cuhk.edu.hk/~king

©2012 Irwin King. All rights reserved.

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
Information and more Information!

flickr  
Amazon.com  
Netflix  
del.icio.us

Wikipedia  
Ebay  
Facebook  
Hulu  
YouTube  
Twitter  
Myspace.com
Consumer Satisfaction  
Company Profit
Real Life Examples
Real Life Examples
Real Life Examples

Yahoo! MOVIES

My Movies: gabe_ma Edit Profile

Recommendations For You

Receive Recommendations by Email

Movies in Theaters: 94089

Burn After Reading (R)
Showtimes & Tickets | Add to My Lists
Yahoo! Users: B- 4794 ratings
The Critics: B 14 reviews

Don't Recommend Again  Seen It? Rate It!

Fight Club (R)
Showtimes & Tickets | Add to My Lists
Yahoo! Users: B+ 52392 ratings
The Critics: B 12 reviews

Don't Recommend Again  Seen It? Rate It!

Vicky Cristina Barcelona (PG-13)
Showtimes & Tickets | Add to My Lists
Yahoo! Users: B 1923 ratings
The Critics: B+ 13 reviews

Don't Recommend Again  Seen It? Rate It!

Pride and Glory (R)
Showtimes & Tickets | Add to My Lists
Yahoo! Users: A- 59 ratings
The Critics: C+ 6 reviews

Don't Recommend Again  Seen It? Rate It!

Lakeview Terrace (PG-13)
Showtimes & Tickets | Add to My Lists
Yahoo! Users: B 3229 ratings
The Critics: C 12 reviews

Don't Recommend Again  Seen It? Rate It!

The Duchess (PG-13)
Showtimes & Tickets | Add to My Lists
Yahoo! Users: B+ 953 ratings
The Critics: B- 10 reviews

Don't Recommend Again  Seen It? Rate It!

See All Recommendations

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
On The Menu

• Introduction
• Basic Techniques
  • Collaborative filtering
  • Matrix factorization
• Different Models
  • Social graph
  • Social ensemble
  • Social distrust
• Website recommendation
Basic Approaches

• Content-based Filtering
  • Recommend items based on key-words
  • More appropriate for information retrieval

• Collaborative Filtering (CF)
  • Look at users with similar rating styles
  • Look at similar items for each item

Underlying assumption: personal tastes are correlated--
Active users will prefer those items which the similar users prefer!
### The tasks

- **Find the unknown rating!**
- **Which item(s) should be recommended?**

---

**Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012**
User-User Similarity

Q1: How to measure the similarity?

Q2: How to select neighbors?
### User-based Collaborative Filtering

<table>
<thead>
<tr>
<th>Users</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>u₁</td>
<td>1</td>
</tr>
<tr>
<td>u₂</td>
<td>3</td>
</tr>
<tr>
<td>u₃</td>
<td></td>
</tr>
<tr>
<td>u₄</td>
<td>1</td>
</tr>
<tr>
<td>u₅</td>
<td></td>
</tr>
<tr>
<td>u₆</td>
<td>1</td>
</tr>
</tbody>
</table>
User-based Collaborative Filtering

<table>
<thead>
<tr>
<th>Users</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>u₁</td>
<td></td>
</tr>
<tr>
<td>u₂</td>
<td>1 3 4 2 5 3 4</td>
</tr>
<tr>
<td>u₃</td>
<td>3 4 3 4 3 4 4</td>
</tr>
<tr>
<td>u₄</td>
<td></td>
</tr>
<tr>
<td>u₅</td>
<td></td>
</tr>
<tr>
<td>u₆</td>
<td>1 3 5 2 4 1 3</td>
</tr>
</tbody>
</table>

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
## User-based Collaborative Filtering

### Items

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Users

<table>
<thead>
<tr>
<th>u1</th>
<th>u2</th>
<th>u3</th>
<th>u4</th>
<th>u5</th>
<th>u6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
### User-based Collaborative Filtering

#### Users

<table>
<thead>
<tr>
<th></th>
<th>u1</th>
<th>u2</th>
<th>u3</th>
<th>u4</th>
<th>u5</th>
<th>u6</th>
</tr>
</thead>
<tbody>
<tr>
<td>u1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>u3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>u5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u6</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

#### Items

- 1
- 3
- 4
- 2
- 5
- 3
- 4
- 4
- 4
- 3
- 5
- 2
- 4
- 1
- 3
**User-based Collaborative Filtering**

<table>
<thead>
<tr>
<th>Users</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_1$</td>
<td>1</td>
</tr>
<tr>
<td>$u_2$</td>
<td>1</td>
</tr>
<tr>
<td>$u_3$</td>
<td></td>
</tr>
<tr>
<td>$u_4$</td>
<td>3</td>
</tr>
<tr>
<td>$u_5$</td>
<td></td>
</tr>
<tr>
<td>$u_6$</td>
<td>1</td>
</tr>
</tbody>
</table>
User-based Collaborative Filtering

• Predict the ratings of active users based on the ratings of similar users found in the user-item matrix

• Pearson correlation coefficient

\[
w(a, i) = \frac{\sum_j (r_{aj} - \bar{r}_a)(r_{ij} - \bar{r}_i)}{\sqrt{\sum_j (r_{aj} - \bar{r}_a)^2 \sum_j (r_{ij} - \bar{r}_i)^2}} \quad j \in I(a) \cap I(i)
\]

• Cosine measure

\[
c(a, i) = \frac{r_a \cdot r_i}{\|r_a\|_2 \ast \|r_i\|_2}
\]
Nearest Neighbor Approaches [Sarwar, 00a]

- Identify highly similar users to the active one
- All with a measure greater than a threshold
- Best K ones

Prediction: \( r_{a,j} = \bar{r}_a + \frac{\sum_i w(a, i)(r_{ij} - \bar{r}_i)}{\sum_i w(a, i)} \)

Figure 1: Three main parts of a Recommender System.
Collaborative Filtering

- **Memory-based Method** (Simple)
  - User-based Method [Xue et al., SIGIR '05]
  - Item-based [Deshpande et al., TOIS '04]
- **Model-based** (Robust)
  - Clustering Methods [Hkors et al, CIMCA '99]
  - Bayesian Methods [Chien et al., IWAIS '99]
  - Aspect Method [Hofmann, SIFIR '03]
  - Matrix Factorization [Sarwar et al., WWW '01]
Collaborative Filtering

• Memory-based (Neighborhood-based)
  • User-based
  • Item-based

• Model-based
  • Clustering Methods
  • Bayesian Methods
  • Matrix Factorization
  • etc.
Collaborative Filtering

- Memory-based (Neighborhood-based)
  - User-based
  - Item-based
- Model-based
  - Clustering Methods
  - Bayesian Methods
  - Matrix Factorization
  - etc...
Matrix Factorization

\[ U = \begin{bmatrix}
1.55 & 1.22 & 0.37 & 0.81 & 0.62 & -0.01 \\
0.36 & 0.91 & 1.21 & 0.39 & 1.10 & 0.25 \\
0.59 & 0.20 & 0.14 & 0.83 & 0.27 & 1.51 \\
0.39 & 1.33 & -0.43 & 0.70 & -0.90 & 0.68 \\
1.05 & 0.11 & 0.17 & 1.18 & 1.81 & 0.40
\end{bmatrix} \quad V = \begin{bmatrix}
1.00 & -0.05 & -0.24 & 0.26 & 1.28 & 0.54 & -0.31 & 0.52 \\
0.19 & -0.86 & -0.72 & 0.05 & 0.68 & 0.02 & -0.61 & 0.70 \\
0.49 & 0.09 & -0.05 & -0.62 & 0.12 & 0.08 & 0.02 & 1.60 \\
-0.40 & 0.70 & 0.27 & -0.27 & 0.99 & 0.44 & 0.39 & 0.74 \\
1.49 & -1.00 & 0.06 & 0.05 & 0.23 & 0.01 & -0.36 & 0.80
\end{bmatrix} \]
Matrix Factorization

- Matrix Factorization in Collaborative Filtering
  - To fit the product of two (low rank) matrices to the observed rating matrix
  - To find two latent user and item feature matrices
  - To use the fitted matrix to predict the unobserved ratings

\[
\begin{pmatrix}
    u_{11} & \cdots & u_{1k} \\
    \vdots & \ddots & \vdots \\
    u_{m1} & \cdots & u_{mk}
\end{pmatrix}
\begin{pmatrix}
    v_{11} & \cdots & v_{1n} \\
    \vdots & \ddots & \vdots \\
    v_{k1} & \cdots & v_{kn}
\end{pmatrix}
\]

- User-specific latent feature vector
- Item-specific latent feature column vector
Matrix Factorization

- Optimization Problem
  - Given a $m \times n$ rating matrix $R$, to find two matrices $U \in \mathbb{R}^{l \times m}$ and $V \in \mathbb{R}^{l \times n}$,

\[
R \approx U^T V,
\]

where $l < \min(m, n)$, is the number of factors
Matrix Factorization

• Models
  • SVD-like Algorithm
  • Regularized Matrix Factorization (RMF)
  • Probabilistic Matrix Factorization (PMF)
  • Non-negative Matrix Factorization (NMF)
SVD-like Algorithm

• Minimizing

\[
\frac{1}{2} \| R - U^T V \|_F^2,
\]

• For collaborative filtering

\[
\min_{U,V} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij} (R_{ij} - U_i^T V_j)^2
\]

where \( I_{ij} \) is the indicator function that is equal to 1 if user \( u_i \) rated item \( v_j \) and equal to 0 otherwise.
Regularized Matrix Factorization

- Minimize the loss based on the observed ratings with regularization terms to avoid over-fitting problem

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

where \( \lambda_1, \lambda_2 > 0 \).

- The problem can be solved by simple gradient descent algorithm.
Social Recommendation Using Probabilistic Matrix Factorization

[Ma et al., CIKM2008]
Challenges

- Data sparsity problem
Challenges

My Movie Ratings

The Pursuit of Happyness (PG-13, 1 hr. 57 min.)
Buy DVD | Add to My Lists
Yahoo! Users: B+ 38992 ratings
The Critics: B- 13 reviews
🌟 My Rating: A+

Finding Nemo (G, 1 hr. 40 min.)
Buy DVD | Add to My Lists
Yahoo! Users: B+ 137394 ratings
The Critics: A- 14 reviews
🌟 My Rating: A

My Blueberry Nights (PG-13, 1 hr. 30 min.)
Buy DVD | Add to My Lists
Yahoo! Users: B- 756 ratings
The Critics: B- 7 reviews
🌟 My Rating: A+

Cold Mountain (R, 2 hrs. 35 min.)
Buy DVD | Add to My Lists
Yahoo! Users: B 38986 ratings
The Critics: B+ 10 reviews
🌟 My Rating: B+

The Lord of the Rings: The Fellowship of the Ring
Buy DVD | Add to My Lists
Yahoo! Users: A- 110957 ratings
The Critics: A 15 reviews
🌟 My Rating: A

Shrek 2 (PG, 1 hr. 32 min.)
Buy DVD | Add to My Lists
Yahoo! Users: B+ 150368 ratings
The Critics: B 15 reviews
🌟 My Rating: B
Challenges

- Traditional recommender systems ignore the social connections between users

Recommendations from friends
Problem Definition

Social Trust Graph

User-Item Rating Matrix
User-Item Matrix Factorization

\[
p(R|U, V, \sigma_R^2) = \prod_{i=1}^m \prod_{j=1}^n \left[ \mathcal{N} \left( R_{ij} | g(U_i^T V_j), \sigma_R^2 \right) \right] I_{ij}^R
\]

\[
p(U|\sigma_U^2) = \prod_{i=1}^m \mathcal{N}(U_i|0, \sigma_U^2 I)
\]

\[
p(V|\sigma_V^2) = \prod_{j=1}^n \mathcal{N}(V_j|0, \sigma_V^2 I)
\]

R. Salakhutdinov and A. Mnih (NIPS'08)

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
SoRec

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
SoRec

\[
\mathcal{L}(R, C, U, V, Z) = \\
\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}^R (r_{ij} - g(U_i^T V_j))^2 + \frac{\lambda_C}{2} \sum_{i=1}^{m} \sum_{k=1}^{m} I_{ik}^C (c_{ik}^* - g(U_i^T Z_k))^2 \\
+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2 + \frac{\lambda_Z}{2} \|Z\|_F^2,
\]

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
SoRec

\[
\frac{\partial L}{\partial U_i} = \sum_{j=1}^{n} I_{ij}^R g'(U^T_i V_j)(g(U^T_i V_j) - r_{ij})V_j \\
+ \lambda_C \sum_{j=1}^{m} I_{ik}^C g'(U^T_i Z_k)(g(U^T_i Z_k) - c^*_{ik})Z_k + \lambda_U U_i,
\]

\[
\frac{\partial L}{\partial V_j} = \sum_{i=1}^{m} I_{ij}^R g'(U^T_i V_j)(g(U^T_i V_j) - r_{ij})U_i + \lambda_V V_j,
\]

\[
\frac{\partial L}{\partial Z_k} = \lambda_C \sum_{i=1}^{m} I_{ik}^C g'(U^T_i Z_k)(g(U^T_i Z_k) - c^*_{ik})U_i + \lambda_Z Z_k,
\]
Disadvantages of SoRec

- Lack of interpretability
- Does not reflect the real-world recommendation process
Learning to Recommend with Social Trust Ensemble

[Ma et al., SIGIR2009]
1st Motivation

- Users have their **own characteristics**, and they have different tastes on different items, such as movies, books, music, articles, food, etc.
2nd Motivation

- Users can be easily influenced by the friends they trust, and prefer their friends’ recommendations.

Where to have dinner?

- Ask
  - Good
  - Very Good
  - Cheap & Delicious
Motivations

• Users have their own characteristics, and they have different tastes on different items, such as movies, books, music, articles, food, etc.

• Users can be easily influenced by the friends they trust, and prefer their friends’ recommendations.

• One user’s final decision is the balance between his/her own taste and his/her trusted friends’ favors.
## User-Item Matrix Factorization

<table>
<thead>
<tr>
<th></th>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$v_3$</th>
<th>$v_4$</th>
<th>$v_5$</th>
<th>$v_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_1$</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_2$</td>
<td>4</td>
<td></td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_3$</td>
<td></td>
<td>2</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_4$</td>
<td>5</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_5$</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

$$p(R|U, V, \sigma^2_R) = \prod_{i=1}^{m} \prod_{j=1}^{n} \left[ \mathcal{N}\left( R_{ij} | g(U_i^T V_j), \sigma^2_R \right) \right]^{I_{ij}}$$

$$p(U | \sigma^2_U) = \prod_{i=1}^{m} \mathcal{N}(U_i | 0, \sigma^2_U \mathbf{I})$$

$$p(V | \sigma^2_V) = \prod_{j=1}^{n} \mathcal{N}(V_j | 0, \sigma^2_V \mathbf{I})$$

[R. Salakhutdinov, et al., NIPS2008]

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
Recommendations by Trusted Friends

\[ \hat{R}_{ik} = \frac{\sum_{j \in T(i)} R_{jk} S_{ij}}{|T(i)|} \]

\[ \hat{R}_{ik} = \sum_{j \in T(i)} R_{jk} S_{ij} \]

\[ p(R|S, U, V, \sigma^2_R) = \prod_{i=1}^{m} \prod_{j=1}^{n} \mathcal{N} \left( R_{ij} | g \left( \sum_{k \in T(i)} S_{ik} U_k^T V_j \right), \sigma^2_S \right)^{I_{ij}^R} \]
Recommendation with Social Trust Ensemble

\[
\prod_{i=1}^{m} \prod_{j=1}^{n} \mathcal{N}\left( R_{ij} \mid g\left( \alpha U_i^T V_j + (1 - \alpha) \sum_{k \in T(i)} S_{ik} U_k^T V_j \right), \sigma^2 \right) \left[ I_{R_{ij}}^R \right]
\]

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
Recommendation with Social Trust Ensemble

\[ \mathcal{L}(R, S, U, V) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} I_{ij}^{R}(R_{ij} - g(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in T(i)} S_{ik} U_k^T V_j))^2 \]
\[ + \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2, \]  

(15)

\[ \frac{\partial \mathcal{L}}{\partial U_i} = \alpha \sum_{j=1}^{n} I_{ij}^{R} g'(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in T(i)} S_{ik} U_k^T V_j) V_j \]
\[ \times (g(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in T(i)} S_{ik} U_k^T V_j) - R_{ij}) \]
\[ + (1 - \alpha) \sum_{p \in B(i)} \sum_{j=1}^{n} I_{pj}^{R} g'(\alpha U_p^T V_j + (1 - \alpha) \sum_{k \in T(p)} S_{pk} U_k^T V_j) \]
\[ \times (g(\alpha U_p^T V_j + (1 - \alpha) \sum_{k \in T(p)} S_{pk} U_k^T V_j) - R_{pj}) S_p V_j \]
\[ + \lambda_U U_i, \]

\[ \frac{\partial \mathcal{L}}{\partial V_j} = \sum_{i=1}^{m} I_{ij}^{R} g'(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in T(i)} S_{ik} U_k^T V_j) \]
\[ \times (g(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in T(i)} S_{ik} U_k^T V_j) - R_{ij}) \]
\[ \times (\alpha U_i + (1 - \alpha) \sum_{k \in T(i)} S_{ik} U_k^T) + \lambda_V V_j, \]
Recommend with Social Distrust

[Ma et al., RecSys2009]
Trust vs. Social

- **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

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
Distrust

• Users’ distrust relations can be interpreted as the “dissimilar” relations

  • On the web, user $U_i$ distrusts user $U_d$ indicates that user $U_i$ disagrees with most of the opinions issued by user $U_d$.

• What to do if a user distrusts many people?

• What to do if many people distrust a user?
Distrust

\[
\max_U \frac{1}{2} \sum_{i=1}^m \sum_{d \in \mathcal{D}^+(i)} S_{id}^D \|U_i - U_d\|_F^2
\]

\[
\min_{U,V} \mathcal{L}_D(R, S^D, U, V) = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (R_{ij} - g(U_i^T V_j))^2 \\
+ \frac{\beta}{2} \sum_{i=1}^m \sum_{d \in \mathcal{D}^+(i)} (-S_{id}^D \|U_i - U_d\|_F^2)
\]

\[
+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2.
\]
Trust

- Users’ trust relations can be interpreted as the “similar” relations
  - On the web, user $U_i$ trusts user $U_t$ indicates that user $U_i$ agrees with most of the opinions issued by user $U_t$. 
Trust

$$\min_U \frac{1}{2} \sum_{i=1}^{m} \sum_{t \in T+(i)} S_{it}^T \|U_i - U_t\|_F^2$$

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

$$+ \frac{\alpha}{2} \sum_{i=1}^{m} \sum_{t \in T+(i)} (S_{it}^T \|U_i - U_t\|_F^2)$$

$$+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2.$$
Web Site Recommendation

[Ma et al., SIGIR 2011]
Traditional Search Paradigm
“Search” to “Discovery”
Challenges in Web Site Recommendation

- Infeasible to ask Web users to explicitly rate Web site

- Not all the traditional methods can be directly applied to the Web site recommendation task

- Can only take advantages of implicit user behavior data
Motivations

- A Web user’s preference can be represented by how frequently a user visits each site.

- Higher visiting frequency on a site means heavy information needs while lower frequency indicates less interests.

- User-query issuing frequency data can be used to refine a user’s preference.
# Using Clicks as Ratings

<table>
<thead>
<tr>
<th>ID</th>
<th>Query</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>358</td>
<td>facebook</td>
<td><a href="http://www.facebook.com">http://www.facebook.com</a></td>
</tr>
<tr>
<td>358</td>
<td>rww</td>
<td><a href="http://www.readthewriteweb.com">http://www.readthewriteweb.com</a></td>
</tr>
<tr>
<td>3968</td>
<td>iphone4</td>
<td><a href="http://www.apple.com">http://www.apple.com</a></td>
</tr>
<tr>
<td>3968</td>
<td>ipad</td>
<td><a href="http://www.apple.com">http://www.apple.com</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Web users</th>
<th>Web sites</th>
<th>Queries</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_1$</td>
<td>$v_1$: 68</td>
<td>$z_1$: 12</td>
</tr>
<tr>
<td>$u_2$</td>
<td>$v_2$: 42</td>
<td>$z_2$: 23</td>
</tr>
<tr>
<td>$u_3$</td>
<td>$v_3$: 72</td>
<td>$z_3$: 14</td>
</tr>
<tr>
<td>$u_4$</td>
<td>$v_4$: 15</td>
<td>$z_4$: 25</td>
</tr>
<tr>
<td>$u_5$</td>
<td>$v_5$: 85</td>
<td>$z_5$: 12</td>
</tr>
<tr>
<td></td>
<td>$v_6$: 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$v_7$: 13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$v_8$: 11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$v_9$: 33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$v_{10}$:</td>
<td></td>
</tr>
</tbody>
</table>
Probabilistic Factor Model

1. Generate $u_{ik} \sim \text{Gamma}(\alpha_k, \beta_k), \forall k$.

2. Generate $v_{jk} \sim \text{Gamma}(\alpha_k, \beta_k), \forall k$.

3. Generate $y_{ij}$ occurrences of item or event $j$ from user $i$ with outcome $y_{ij} = \sum_{k=1}^{d} u_{ik} v_{jk}$.

4. Generate $f_{ij} \sim \text{Poisson}(y_{ij})$.

\[
p(U|\alpha, \beta) = \prod_{i=1}^{m} \prod_{k=1}^{d} \frac{u_{ik}^{\alpha_k-1} \exp(-u_{ik}/\beta_k)}{\beta_k^{\alpha_k} \Gamma(\alpha_k)}
\]

\[
p(V|\alpha, \beta) = \prod_{j=1}^{n} \prod_{k=1}^{d} \frac{v_{jk}^{\alpha_k-1} \exp(-v_{jk}/\beta_k)}{\beta_k^{\alpha_k} \Gamma(\alpha_k)}
\]

\[
p(F|Y) = \prod_{i=1}^{m} \prod_{j=1}^{n} \frac{y_{ij}^{f_{ij}} \exp(-y_{ij})}{f_{ij}!}
\]

\[
p(U, V|F, \alpha, \beta) \propto p(F|Y)p(U|\alpha, \beta)p(V|\alpha, \beta)
\]

\[
\mathcal{L}(U, V; F) = \sum_{i=1}^{m} \sum_{k=1}^{d} ((\alpha_k - 1) \ln(u_{ik}/\beta_k) - u_{ik}/\beta_k) + \sum_{j=1}^{n} \sum_{k=1}^{d} ((\alpha_k - 1) \ln(v_{jk}/\beta_k) - v_{jk}/\beta_k) + \sum_{i=1}^{m} \sum_{j=1}^{n} (f_{ij} \ln y_{ij} - y_{ij}) + \text{const.}
\]
Probabilistic Factor Model

\[ \mathcal{L}(U, V; F) = \sum_{i=1}^{m} \sum_{k=1}^{d} \left( \alpha_k - 1 \right) \ln \left( u_{ik} / \beta_k \right) - u_{ik} / \beta_k \]

\[ + \sum_{j=1}^{n} \sum_{k=1}^{d} \left( \alpha_k - 1 \right) \ln \left( v_{jk} / \beta_k \right) - v_{jk} / \beta_k \]

\[ + \sum_{i=1}^{m} \sum_{j=1}^{n} \left( f_{ij} \ln y_{ij} - y_{ij} \right) + \text{const.} \]

\[ u_{ik} \leftarrow u_{ik} \frac{\sum_{j=1}^{n} \left( f_{ij} v_{jk} / y_{ij} \right) + \left( \alpha_k - 1 \right) / u_{ik}}{\sum_{j=1}^{n} v_{jk} + 1 / \beta_k} \]

\[ v_{jk} \leftarrow v_{jk} \frac{\sum_{i=1}^{m} \left( f_{ij} u_{ik} / y_{ij} \right) + \left( \alpha_k - 1 \right) / v_{jk}}{\sum_{i=1}^{m} u_{ik} + 1 / \beta_k} \]
Collective Probabilistic Factor Model

\[ \mathcal{L}(U, V, Z; F^x, F^y) \]

\[
= \sum_{i=1}^{m} \sum_{l=1}^{p} (f^x_{il} \ln x_{il} - x_{il}) + \sum_{i=1}^{m} \sum_{j=1}^{n} (f^y_{ij} \ln y_{ij} - y_{ij}) \\
+ \sum_{i=1}^{m} \sum_{k=1}^{d} ((\alpha_k - 1) \ln (u_{ik} / \beta_k) - u_{ik} / \beta_k) \\
+ \sum_{j=1}^{n} \sum_{k=1}^{d} ((\alpha_k - 1) \ln (v_{jk} / \beta_k) - v_{jk} / \beta_k) \\
+ \sum_{i=1}^{m} \sum_{k=1}^{d} ((\alpha_k - 1) \ln (z_{lk} / \beta_k) - z_{lk} / \beta_k) + \text{const.}
\]

\[
u_{ik} \leftarrow u_{ik} \frac{\sum_{j=1}^{n} f^y_{ij} v_{jk} / y_{ij} + (\alpha_k - 1) / v_{jk}}{\sum_{j=1}^{n} v_{jk} + \sum_{l=1}^{p} z_{lk} + 1 / \beta_k},
\]

\[
v_{jk} \leftarrow v_{jk} \frac{\sum_{i=1}^{m} (f^y_{ij} u_{ik} / y_{ij}) + (\alpha_k - 1) / v_{jk}}{\sum_{i=1}^{m} u_{ik} + 1 / \beta_k},
\]

\[
z_{lk} \leftarrow z_{lk} \frac{\sum_{i=1}^{m} (f^x_{il} u_{ik} / x_{il}) + (\alpha_k - 1) / z_{lk}}{\sum_{i=1}^{m} u_{ik} + 1 / \beta_k}.
\]

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
Dataset

- Anonymous logs of Web sites visited by users who opted-in to provide data through browser toolbar
- URLs of all the Web sites are truncated to the site level
- After pruning one month data, we have 165,403 users, 265,367 URLs and 442,598 queries
- User-site frequency matrix has 2,612,016 entries, while in user-query frequency matrix has 833,581 entries

Table 2: Statistics of User-Site and User-Query Frequency Matrices

<table>
<thead>
<tr>
<th>Statistics</th>
<th>User-Site Frequency</th>
<th>User-Query Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min. Num.</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Max. Num.</td>
<td>9,969</td>
<td>4,693</td>
</tr>
<tr>
<td>Avg. Num.</td>
<td>20.33</td>
<td>23.05</td>
</tr>
</tbody>
</table>
## Performance Comparison

**Table 3: Performance Comparison (Dimensionality = 10)**

<table>
<thead>
<tr>
<th>Training Data</th>
<th>Metrics</th>
<th>UserMean</th>
<th>SiteMean</th>
<th>SVD</th>
<th>PMF</th>
<th>NMF</th>
<th>GaP</th>
<th>PFM</th>
<th>CPFM</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>NMAE</td>
<td>2.246</td>
<td>1.094</td>
<td>0.488</td>
<td>0.476</td>
<td>0.465</td>
<td>0.440</td>
<td>0.432</td>
<td>0.427</td>
</tr>
<tr>
<td></td>
<td>Improve</td>
<td>80.98%</td>
<td>60.96%</td>
<td>12.50%</td>
<td>10.29%</td>
<td>8.17%</td>
<td>2.95%</td>
<td>0.529</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>NRMSE</td>
<td>3.522</td>
<td>2.171</td>
<td>0.581</td>
<td>0.570</td>
<td>0.554</td>
<td>0.532</td>
<td>0.529</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>Improve</td>
<td>85.24%</td>
<td>76.05%</td>
<td>10.50%</td>
<td>8.77%</td>
<td>6.14%</td>
<td>2.26%</td>
<td>0.434</td>
<td>0.428</td>
</tr>
<tr>
<td>80%</td>
<td>NMAE</td>
<td>2.252</td>
<td>1.096</td>
<td>0.490</td>
<td>0.478</td>
<td>0.468</td>
<td>0.441</td>
<td>0.434</td>
<td>0.428</td>
</tr>
<tr>
<td></td>
<td>Improve</td>
<td>80.99%</td>
<td>60.95%</td>
<td>12.65%</td>
<td>10.46%</td>
<td>8.55%</td>
<td>2.95%</td>
<td>0.434</td>
<td>0.428</td>
</tr>
<tr>
<td></td>
<td>NRMSE</td>
<td>3.714</td>
<td>2.159</td>
<td>0.584</td>
<td>0.571</td>
<td>0.560</td>
<td>0.533</td>
<td>0.530</td>
<td>0.520</td>
</tr>
<tr>
<td></td>
<td>Improve</td>
<td>86.00%</td>
<td>75.91%</td>
<td>10.96%</td>
<td>8.93%</td>
<td>7.14%</td>
<td>2.44%</td>
<td>0.530</td>
<td>0.520</td>
</tr>
</tbody>
</table>

**Table 4: Performance Comparison (Dimensionality = 20)**

<table>
<thead>
<tr>
<th>Training Data</th>
<th>Metrics</th>
<th>UserMean</th>
<th>SiteMean</th>
<th>SVD</th>
<th>PMF</th>
<th>NMF</th>
<th>GaP</th>
<th>PFM</th>
<th>CPFM</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>NMAE</td>
<td>2.246</td>
<td>1.094</td>
<td>0.469</td>
<td>0.460</td>
<td>0.449</td>
<td>0.426</td>
<td>0.413</td>
<td>0.409</td>
</tr>
<tr>
<td></td>
<td>Improve</td>
<td>81.79%</td>
<td>62.61%</td>
<td>12.79%</td>
<td>11.09%</td>
<td>8.91%</td>
<td>3.99%</td>
<td>0.503</td>
<td>0.496</td>
</tr>
<tr>
<td></td>
<td>NRMSE</td>
<td>3.522</td>
<td>2.171</td>
<td>0.568</td>
<td>0.556</td>
<td>0.542</td>
<td>0.521</td>
<td>0.503</td>
<td>0.496</td>
</tr>
<tr>
<td></td>
<td>Improve</td>
<td>85.92%</td>
<td>77.15%</td>
<td>12.68%</td>
<td>10.79%</td>
<td>8.49%</td>
<td>4.80%</td>
<td>0.503</td>
<td>0.496</td>
</tr>
<tr>
<td>80%</td>
<td>NMAE</td>
<td>2.252</td>
<td>1.096</td>
<td>0.470</td>
<td>0.462</td>
<td>0.451</td>
<td>0.427</td>
<td>0.415</td>
<td>0.410</td>
</tr>
<tr>
<td></td>
<td>Improve</td>
<td>81.79%</td>
<td>62.59%</td>
<td>12.77%</td>
<td>11.26%</td>
<td>9.09%</td>
<td>3.98%</td>
<td>0.415</td>
<td>0.410</td>
</tr>
<tr>
<td></td>
<td>NRMSE</td>
<td>3.714</td>
<td>2.159</td>
<td>0.570</td>
<td>0.558</td>
<td>0.545</td>
<td>0.522</td>
<td>0.504</td>
<td>0.498</td>
</tr>
<tr>
<td></td>
<td>Improve</td>
<td>86.59%</td>
<td>76.93%</td>
<td>12.63%</td>
<td>10.75%</td>
<td>8.62%</td>
<td>4.60%</td>
<td>0.504</td>
<td>0.498</td>
</tr>
</tbody>
</table>
Impact of Parameters

Figure 6: Impact of Parameter $\alpha_k$ in PFM

(a) NMAE

(b) NRMSE

(c) Gamma Distributions

Figure 7: Impact of Parameter $\beta_k$ in PFM

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
Impact of Parameters

Figure 8: Impact of Parameter $\theta$ in CPFM
Concluding Remarks

- Social recommendation extends traditional models and techniques by using social graphs, ensembles, distrust relationships, clicks, etc.

- Fusing of social behavior information, e.g., social relationships, personal preferences, media consumption patterns, temporal dynamics, location information, etc. provides better models for social recommendations.
Acknowledgments

- Ritesh Agrawal
- David Grangier
- Jean-Francois Paiement
- Remi Zajac
- Shouyuan Chen (Ph.D.)
- Baichuan Li (Ph.D.)
- Zhenjiang Lin (Ph.D.)
- Guang Ling (Ph.D.)
- Hao Ma (Microsoft)
- Haiqin Yang (Postdoc)
- Connie Yuen (Ph.D.)
- Xin Xin (Postdoc)
- Chao Zhou (Ph.D.)
On-Going Research

Machine Learning

- Smooth Optimization for Effective Multiple Kernel Learning (AAAI’10)
- Simple and Efficient Multiple Kernel Learning By Group Lasso (ICML’10)
- Online Learning for Group Lasso (ICML’10)
- Heavy-Tailed Symmetric Stochastic Neighbor Embedding (NIPS’09)
- Adaptive Regularization for Transductive Support Vector Machine (NIPS’09)
- Direct Zero-norm Optimization for Feature Selection (ICDM’08)
- Semi-supervised Learning from General Unlabeled Data (ICDM’08)
- Learning with Consistency between Inductive Functions and Kernels (NIPS’08)
- An Extended Level Method for Efficient Multiple Kernel Learning (NIPS’08)
- Semi-supervised Text Categorization by Active Search (CIKM’08)
- Transductive Support Vector Machine (NIPS’07)
- Global and local learning (ICML’04, JMLR’04)
On-Going Research

Web Intelligence/Information Retrieval

- Learning to Suggest Questions in Online Forums (AAAI’11)
- Diversifying Query Suggestion Results (AAAI’10)
- A Generalized Co-HITS Algorithm and Its Application to Bipartite Graphs (KDD’09)
- Entropy-biased Models for Query Representation on the Click Graph (SIGIR’09)
- Effective Latent Space Graph-based Re-ranking Model with Global Consistency (WSDM’09)
- Formal Models for Expert Finding on DBLP Bibliography Data (ICDM’08)
- Learning Latent Semantic Relations from Query Logs for Query Suggestion (CIKM’08)
- RATE: a Review of Reviewers in a Manuscript Review Process (WI’08)
- MatchSim: link-based web page similarity measurements (WI’07)
- Diffusion rank: Ranking web pages based on heat diffusion equations (SIGIR’07)
- Web text classification (WWW’07)
On-Going Research

**Recommender Systems/Collaborative Filtering**

- Probabilistic Factor Models for Web Site Recommendation (*SIGIR’11*)
- Recommender Systems with Social Regularization (*WSDM’11*)
- UserRec: A User Recommendation Framework in Social Tagging Systems (*AAAI’10*)
- Learning to Recommend with Social Trust Ensemble (*SIRIR’09*)
- Semi-Nonnegative Matrix Factorization with Global Statistical Consistency in Collaborative Filtering (*CIKM’09*)
- Recommender system: accurate recommendation based on sparse matrix (*SIGIR’07*)
- SoRec: Social Recommendation Using Probabilistic Matrix Factorization (*CIKM’08*)

**Human Computation**

- A Survey of Human Computation Systems (*SCA’09*)
- Mathematical Modeling of Social Games (*SIAG’09*)
- An Analytical Study of Puzzle Selection Strategies for the ESP Game (*WI’08*)
- An Analytical Approach to Optimizing The Utility of ESP Games (*WI’08*)
Ever since its inception, the Web has changed the landscape of human experiences on how we interact with one another and data through service infrastructures via various computing devices. This interweaving environment is now becoming ever more embedded into devices and systems that integrate seamlessly on how we live, both in our working or leisure time.

For this volume, King and Baeza-Yates selected some pioneering and cutting-edge research work that is pointing to the future of the Web. Based on the Workshop Track of the 17th International World Wide Web Conference (WWW2008) in Beijing, they selected the top contributions and asked the authors to resubmit their work with a minimum of one third of additional material from their original workshop manuscripts to be considered for this volume. After a second round of reviews and selection, 16 contributions were finally accepted.

The work within this volume represents the tip of an iceberg of the many exciting advancements on the WWW. It covers topics like semantic web services, location-based and mobile applications, personalized and context-dependent user interfaces, social networks, and folksonomies. The presentations aim at researchers in academia and industry by showcasing latest research findings. Overall, they deliver an excellent picture of the current state-of-the-art, and will also serve as the basis for ongoing research discussions and point to new directions.
VeriGuide

- **Similarity text** detection system
- Developed at **CUHK**
- Promote and uphold academic honesty, integrity, and quality
- Support **English, Traditional** and **Simplified Chinese**
- Handle `.doc`, `.txt`, `.pdf`, `.html`, etc. file formats
- Generate detailed **originality report** including readability

Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012
Q & A