

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.



You Tube



ebay

facebook

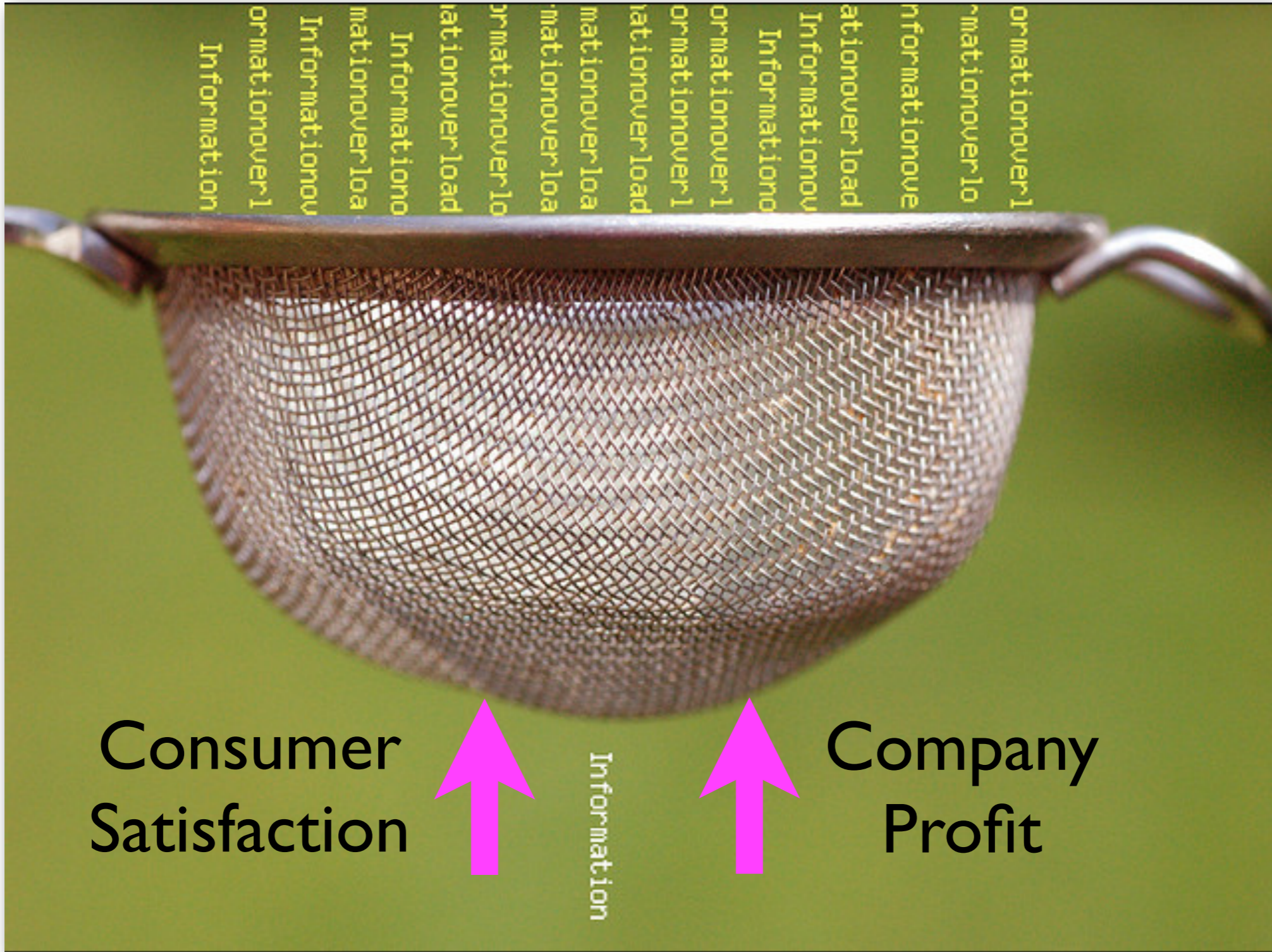
hulu™

twitter



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

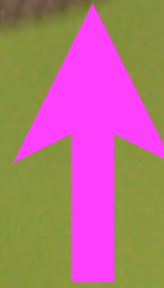




Consumer Satisfaction



Information



Company Profit



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



Real Life Examples

The screenshot shows the Amazon.com interface for a book. At the top, the Amazon logo is on the left, and navigation links like 'Hello. Sign in', 'Your Amazon.com', 'Today's Deals', 'Gifts & Wish Lists', 'Gift Cards', 'Your Digital Items', 'Your Account', and 'Help' are on the right. Below this is a search bar with 'Books' entered and a 'GO' button. A navigation bar below the search bar includes 'Shop All Departments', 'Books', 'Advanced Search', 'Browse Subjects', 'New Releases', 'Bestsellers', 'The New York Times® Bestsellers', 'Libros en español', 'Bargain Books', and 'Textbooks'. The main product area features a book cover on the left with a 'Click to LOOK INSIDE!' callout. The book title is 'Weaving Services and People on the World Wide Web [Hardcover]' by Irwin King (Editor) and Ricardo Baeza-Yates (Editor). The price is \$79.11, down from a list price of \$99.00, with a 20% discount. It is 'In Stock' and ships from Amazon.com. To the right of the book details are purchase options: 'Add to Cart' (quantity 1), 'Add to Cart with FREE Two-Day Shipping' (requires Amazon Prime), and 'Add to Wish List'. Below these are 'More Buying Choices' showing 31 used and new options from \$14.62, and a 'Sell yours here' button. At the bottom of the product area are social sharing icons and a 'FREE Two-Day Shipping for Students' badge.

amazon.com Hello. [Sign in](#) to get personalized recommendations. New customer? [Start here.](#) **FREE 2-Day Shipping: See details**

Your Amazon.com Today's Deals | [Gifts & Wish Lists](#) | [Gift Cards](#) | Your Digital Items | Your Account | Help

Shop All Departments Search Cart Wish List

Books | [Advanced Search](#) | [Browse Subjects](#) | [New Releases](#) | [Bestsellers](#) | [The New York Times® Bestsellers](#) | [Libros en español](#) | [Bargain Books](#) | [Textbooks](#)

Click to **LOOK INSIDE!**

Weaving Services and People on the World Wide Web [Hardcover]
Irwin King (Editor), Ricardo Baeza-Yates (Editor)
[Be the first to review this item](#) | Like (0)

List Price: ~~\$99.00~~
Price: **\$79.11** & this item ships for **FREE with Super Saver Shipping.** [Details](#)
You Save: **\$19.89 (20%)**

In Stock.
Ships from and sold by Amazon.com. Gift-wrap available.
Only 2 left in stock--order soon (more on the way).
Want it delivered Thursday, July 21? Order it in the next 0 hours and 19 minutes, and choose **One-Day Shipping** at checkout. [Details](#)
18 new from \$14.62 **13 used** from \$14.62

FREE Two-Day Shipping for Students. [Learn more](#)

Quantity:

Add to Cart
or
[Sign in](#) to turn on 1-Click ordering.
or
 Add to Cart with FREE Two-Day Shipping
Amazon Prime Free Trial required. Sign up when you check out. [Learn More](#)

Add to Wish List

More Buying Choices
31 used & new from \$14.62
Have one to sell? **Sell yours here**

[Share](#)

[Share your own customer images](#)
[Search inside this book](#)



Customers Who Bought This Item Also Bought



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



Real Life Examples

amazon.com | Hao's Amazon.com | See All 40 Product Categories | Your Account |  Cart | Your Lists | Help | 

Your Browsing History | Recommended For You | Rate These Items | Improve Your Recommendations | Your Profile | Learn More

Today's Recommendations For You

Here's a daily sample of items recommended for you. Click here to [see all recommendations](#).

Page 1 of 25



[Invincible](#) ~ Michael Jackson
★★★★☆ (880) \$7.99



[In Search of Sunrise, Vol. 7: Asia](#) ~ DJ Tiesto
★★★★☆ (53) \$15.99



[Fallen](#) ~ Evanescence
★★★★☆ (2,447) \$8.99



[Amar Es Combatir](#) ~ Maná
★★★★☆ (55) \$8.49



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



Real Life Examples



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!



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!



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!



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!



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!



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

Underling assumption: personal tastes are correlated--
Active users will prefer those items which the similar
users prefer!



Framework

		Items												
		i_1	i_2			i_j							i_m	
Users	u_1													
	u_2	1	3		4		2		5			3	4	
	u_i		3		4		r_{ij}	3	4		3	4		4
	u_n	1			3	5	2		4	1			3	

- The tasks

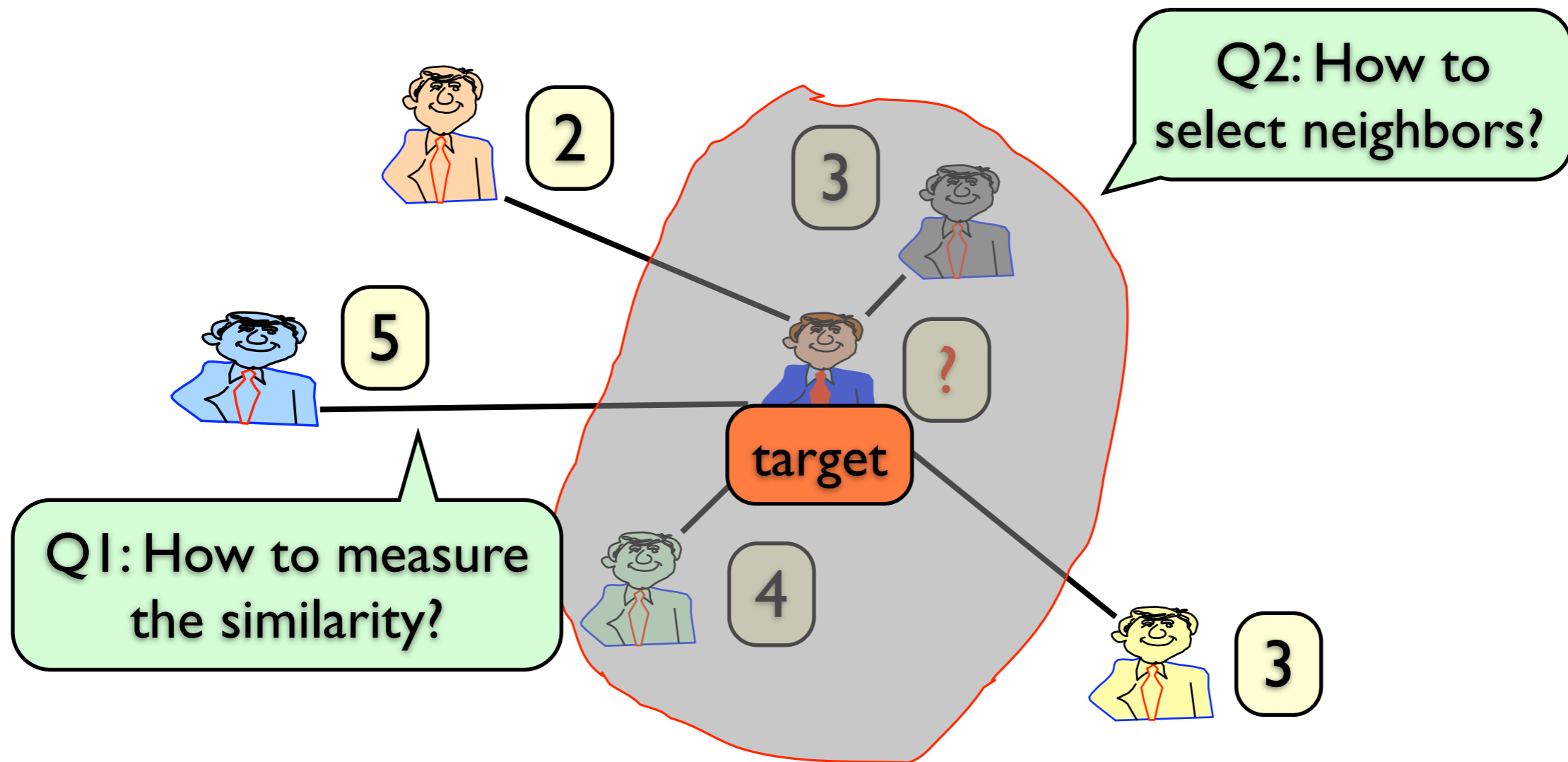
- Find the **unknown** rating!
- Which **item(s)** should be recommended?



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



User-User Similarity



User-based Collaborative Filtering

Items

Users

u ₁													
u ₂	1	3		4		2		5			3	4	
u ₃													
u ₄		3		4			3	4		3	4		4
u ₅													
u ₆	1			3	5	2		4	1			3	



User-based Collaborative Filtering

Items

Users

u ₁												
u ₂	1	3		4		2		5			3	4
u ₃												
u ₄		3		4			3	4		3	4	4
u ₅												
u ₆	1			3	5	2		4	1			3



User-based Collaborative Filtering

Items

Users

u ₁													
u ₂	1	3	4	2	5			3	4				
u ₃													
u ₄		3	4		3	4		3	4		4		
u ₅													
u ₆	1		3	5	2	4	1				3		



User-based Collaborative Filtering

Items

Users

u ₁													
u ₂	1	3		4		2		5			3	4	
u ₃													
u ₄		3		4			3	4		3	4		4
u ₅													
u ₆	1			3	5	2		4	1			3	



User-based Collaborative Filtering

Items

Users

u ₁													
u ₂	1	3	4		2		5			3	4		
u ₃													
u ₄		3	4			3	4		3	4		4	
u ₅													
u ₆	1		3	5	2		4	1			3		



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

$$c(a, i) = \frac{r_a \cdot r_i}{\|r_a\|_2 * \|r_i\|_2}$$

u_i	1	3	4	2	5		3	4				
u_a	3	4		3	4		3	4		4		
	1		3	5	2		4	1		3		



Nearest Neighbor Approaches

[Sarwar, 00a]

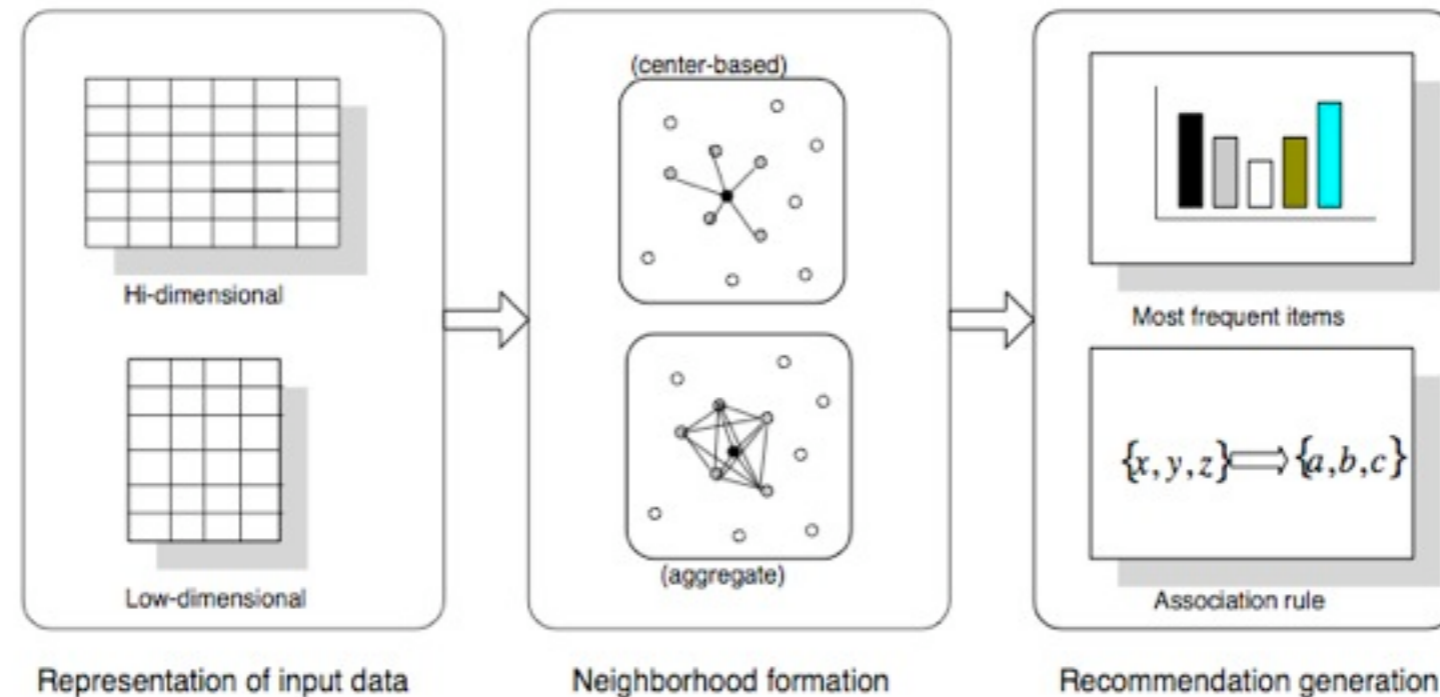


Figure 1: Three main parts of a Recommender System.

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

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



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



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., IW AIS '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

	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8
u_1	5	2		3		4		
u_2	4	3			5			
u_3	4		2				2	4
u_4								
u_5	5	1	2		4	3		
u_6	4	3		2	4		3	5

	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8
u_1	5	2	2.5	3	4.8	4	2.2	4.8
u_2	4	3	2.4	2.9	5	4.1	2.6	4.7
u_3	4	1.7	2	3.2	3.9	3.0	2	4
u_4	4.8	2.1	2.7	2.6	4.7	3.8	2.4	4.9
u_5	5	1	2	3.4	4	3	1.5	4.6
u_6	4	3	2.9	2	4	3.4	3	5

$$U = \begin{bmatrix} 1.55 & 1.22 & 0.37 & 0.81 & 0.62 & -0.01 \\ 0.36 & 0.91 & 1.21 & 0.39 & 1.10 & 0.25 \\ 0.59 & 0.20 & 0.14 & 0.83 & 0.27 & 1.51 \\ 0.39 & 1.33 & -0.43 & 0.70 & -0.90 & 0.68 \\ 1.05 & 0.11 & 0.17 & 1.18 & 1.81 & 0.40 \end{bmatrix}$$

$$V = \begin{bmatrix} 1.00 & -0.05 & -0.24 & 0.26 & 1.28 & 0.54 & -0.31 & 0.52 \\ 0.19 & -0.86 & -0.72 & 0.05 & 0.68 & 0.02 & -0.61 & 0.70 \\ 0.49 & 0.09 & -0.05 & -0.62 & 0.12 & 0.08 & 0.02 & 1.60 \\ -0.40 & 0.70 & 0.27 & -0.27 & 0.99 & 0.44 & 0.39 & 0.74 \\ 1.49 & -1.00 & 0.06 & 0.05 & 0.23 & 0.01 & -0.36 & 0.80 \end{bmatrix}$$



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

$$\begin{pmatrix} \mathbf{u}_{11} & \dots & \mathbf{u}_{1k} \\ \vdots & \ddots & \vdots \\ \mathbf{u}_{m1} & \dots & \mathbf{u}_{mk} \end{pmatrix} \begin{pmatrix} \mathbf{v}_{11} & \dots & \mathbf{v}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{v}_{k1} & \dots & \mathbf{v}_{kn} \end{pmatrix}$$

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

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

Regularization terms

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]



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



Challenges

- Data sparsity problem

YAHOO! MOVIES

My Movies: **gabe_ma** [Edit Profile](#)

Recommendations For You



[Watch the Trailer](#)

My Blueberry Nights (2008)

The Critics:

B-

[7 reviews](#)

My Grade:

A+

Oscar-worthy

A

B

C

D

F

Yahoo! Users:

B-

[667 ratings](#)

[write a review](#)



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!](#)



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



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



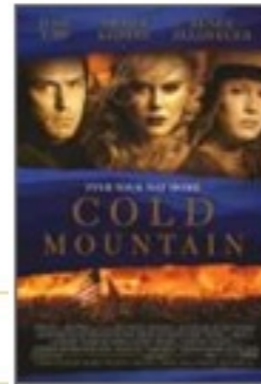
My Blueberry Nights (PG-13, 1 hr. 30 min.)

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

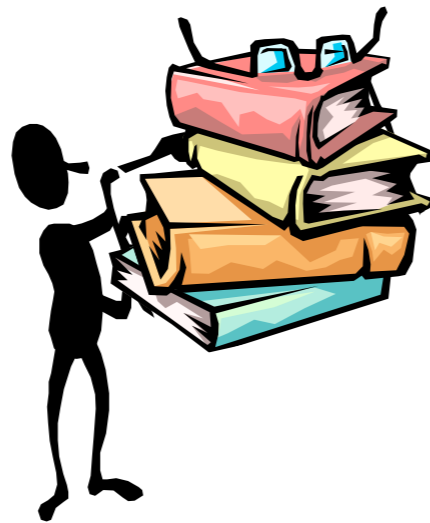


Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012



Challenges

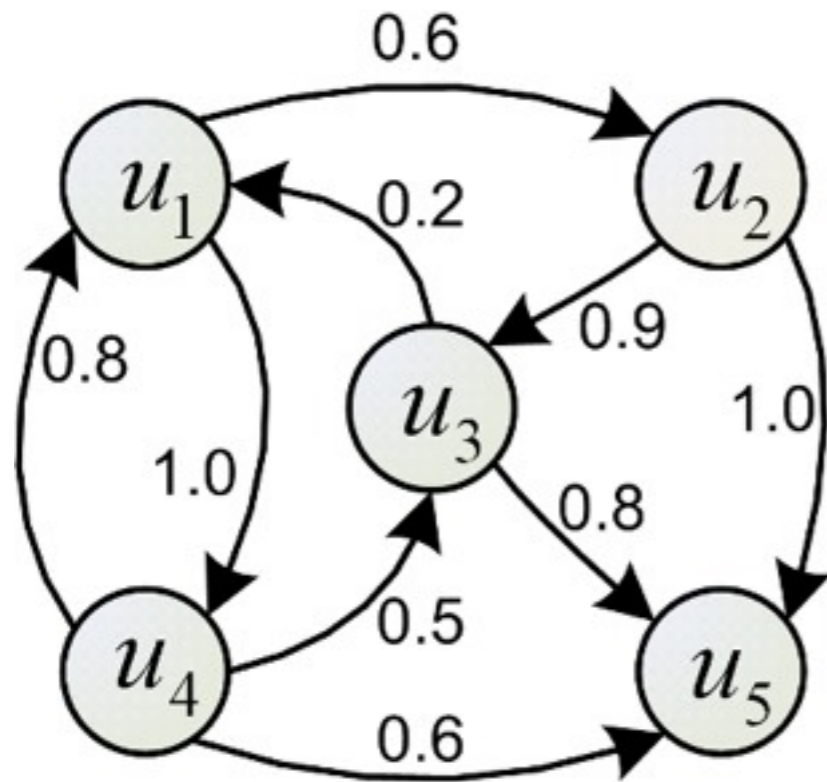
- Traditional recommender systems ignore the social connections between users



Recommendations
from friends



Problem Definition



Social Trust Graph

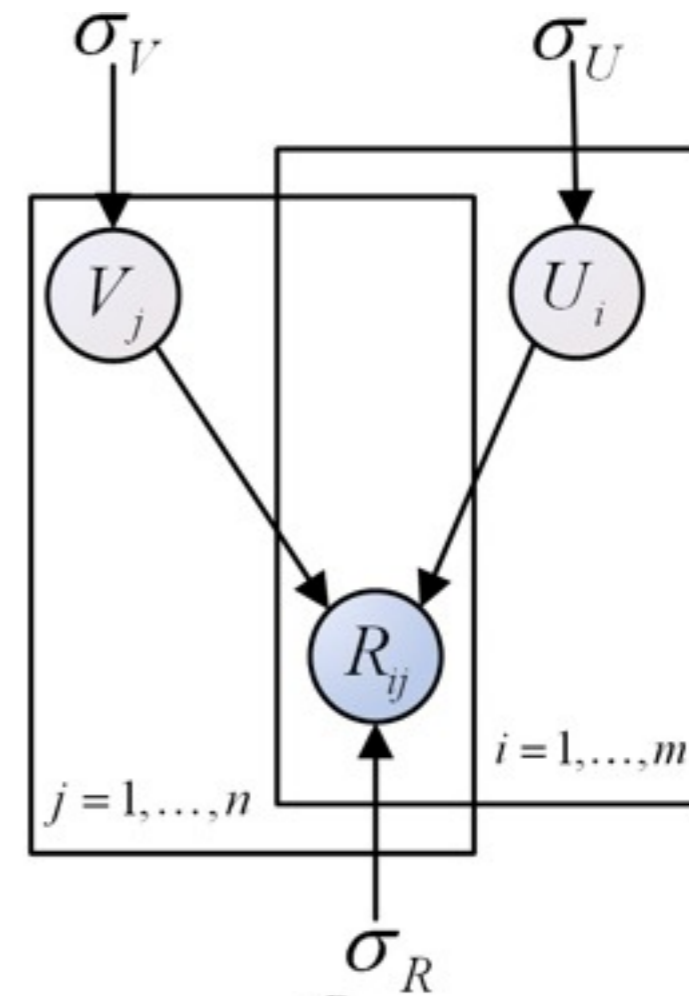
	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3

User-Item Rating Matrix



User-Item Matrix Factorization

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3



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

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

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

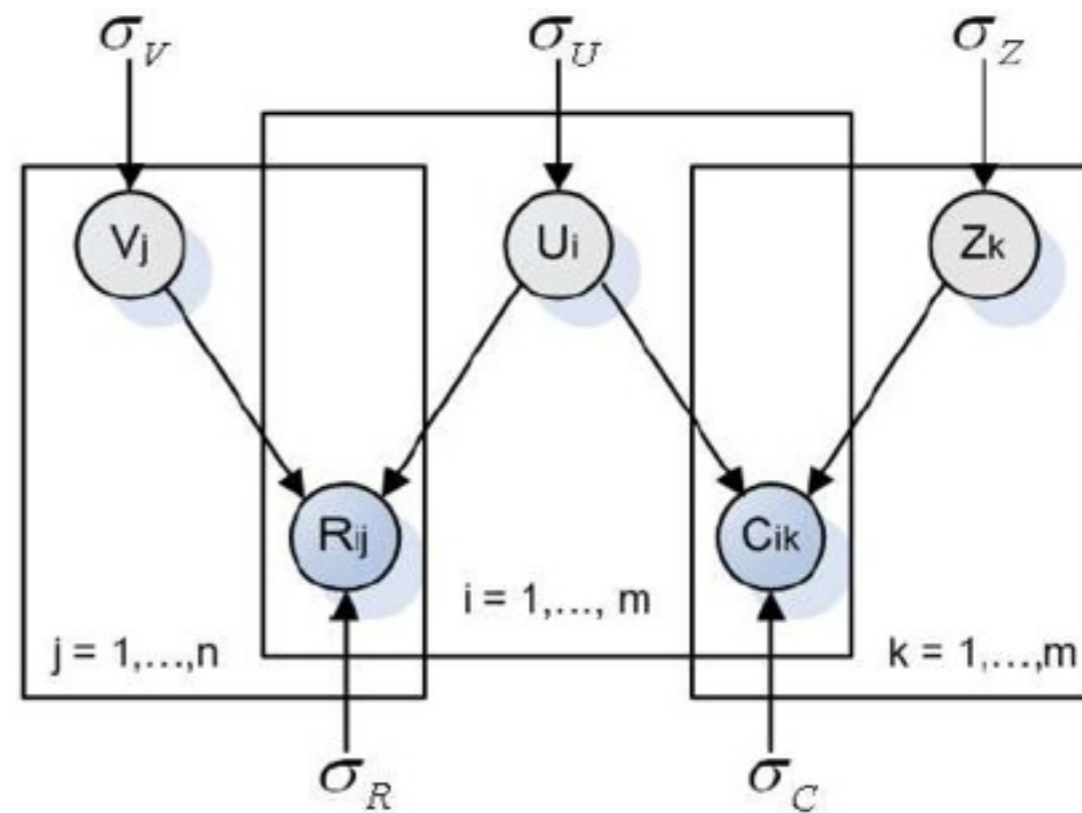
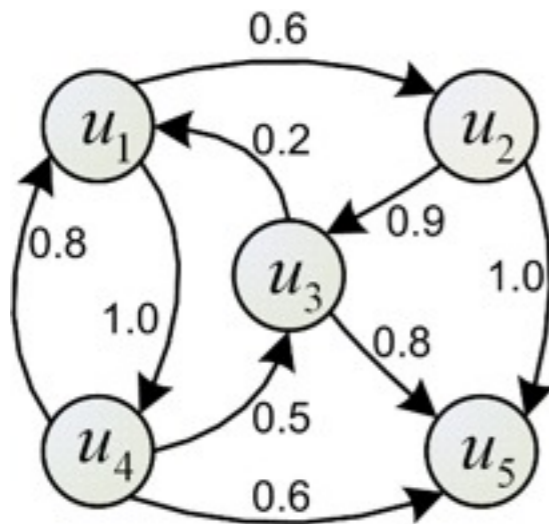
R. Salakhutdinov and A. Mnih (NIPS'08)

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



SoRec

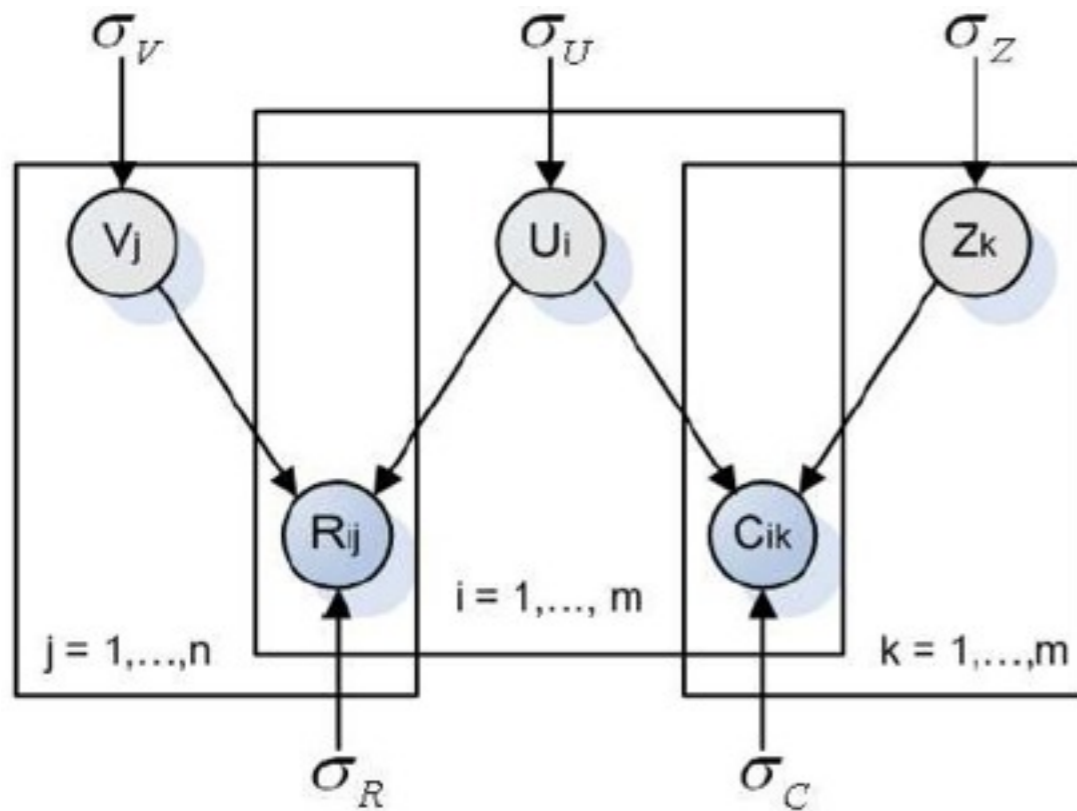
	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3



SoRec



SoRec



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

$$p(C|U, Z, \sigma_C^2) = \prod_{i=1}^m \prod_{k=1}^m \mathcal{N} \left[\left(c_{ik} | g(U_i^T Z_k), \sigma_C^2 \right) \right]^{I_{ik}^C}$$

$$p(U|\sigma_U^2) = \prod_{i=1}^m \mathcal{N}(U_i | 0, \sigma_U^2 \mathbf{I}) \quad p(V|\sigma_V^2) = \prod_{j=1}^n \mathcal{N}(V_j | 0, \sigma_V^2 \mathbf{I})$$

$$p(Z|\sigma_Z^2) = \prod_{k=1}^m \mathcal{N}(Z_k | 0, \sigma_Z^2 \mathbf{I})$$

$$\mathcal{L}(R, C, U, V, Z) =$$

$$\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (r_{ij} - g(U_i^T V_j))^2 + \frac{\lambda_C}{2} \sum_{i=1}^m \sum_{k=1}^m I_{ik}^C (c_{ik}^* - g(U_i^T Z_k))^2$$

$$+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2 + \frac{\lambda_Z}{2} \|Z\|_F^2,$$



SoRec

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial U_i} &= \sum_{j=1}^n I_{ij}^R g'(U_i^T V_j) (g(U_i^T V_j) - r_{ij}) V_j \\ &+ \lambda_C \sum_{k=1}^m I_{ik}^C g'(U_i^T Z_k) (g(U_i^T Z_k) - c_{ik}^*) Z_k + \lambda_U U_i, \end{aligned}$$

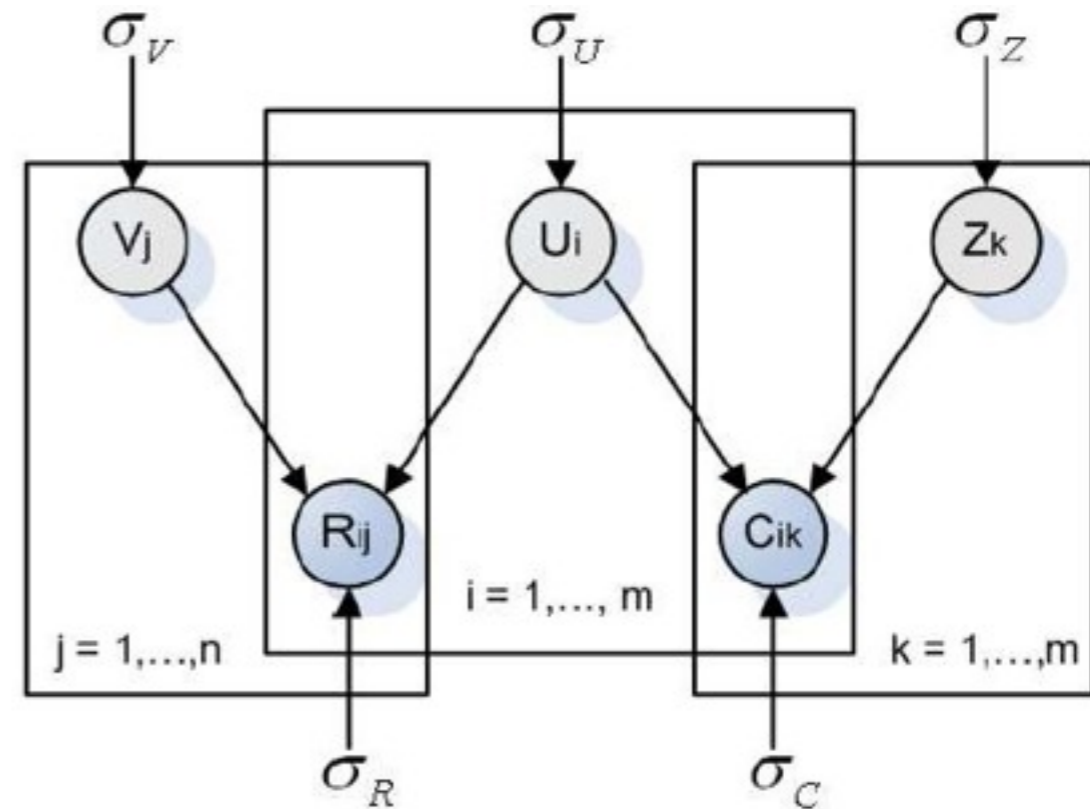
$$\frac{\partial \mathcal{L}}{\partial V_j} = \sum_{i=1}^m I_{ij}^R g'(U_i^T V_j) (g(U_i^T V_j) - r_{ij}) U_i + \lambda_V V_j,$$

$$\frac{\partial \mathcal{L}}{\partial Z_k} = \lambda_C \sum_{i=1}^m I_{ik}^C g'(U_i^T Z_k) (g(U_i^T Z_k) - c_{ik}^*) U_i + \lambda_Z Z_k,$$



Disadvantages of SoRec

- Lack of interpretability
- Does not reflect the real-world recommendation process



SoRec



Learning to Recommend with Social Trust Ensemble

[Ma et al., SIGIR2009]



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



1st Motivation

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

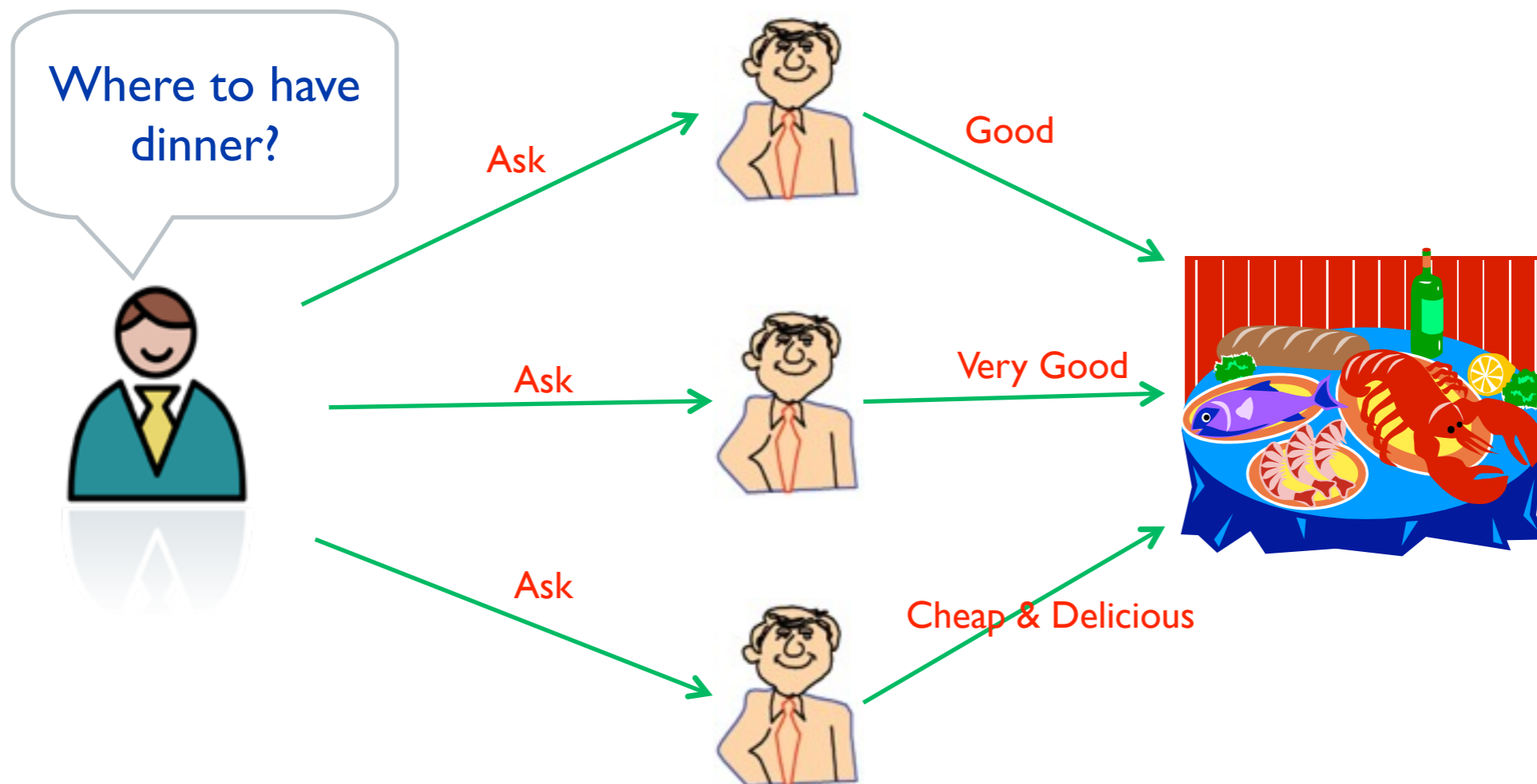


Social Recommendation, Irwin King @ Bay Area Search Group Meetup, February 22, 2012



2nd Motivation

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



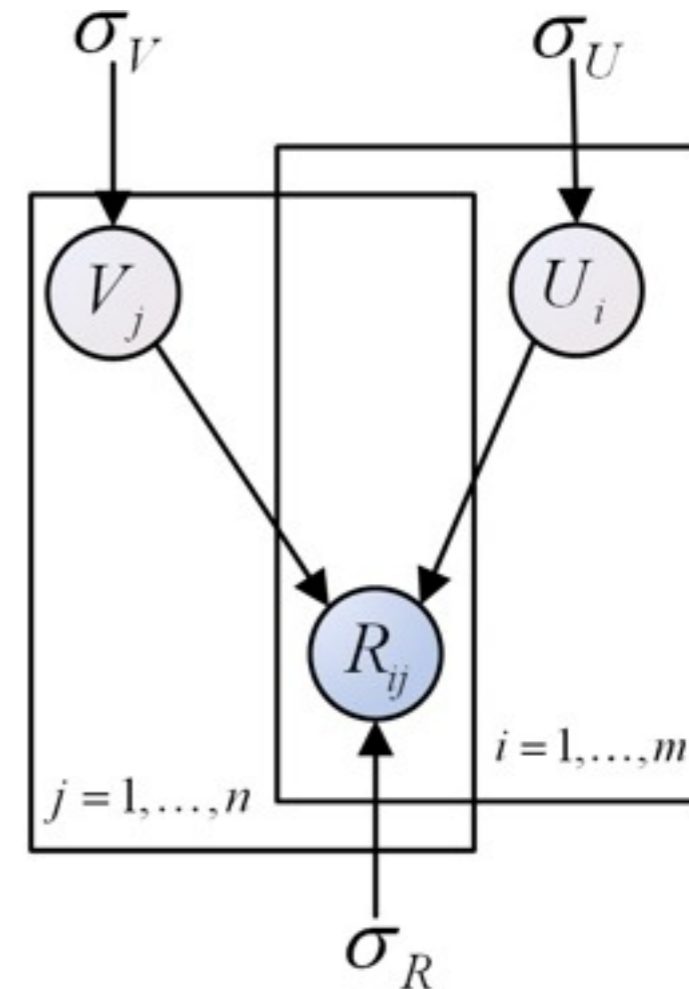
Motivations

- 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

	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3



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

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

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

[R. Salakhutdinov, et al., NIPS2008]

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



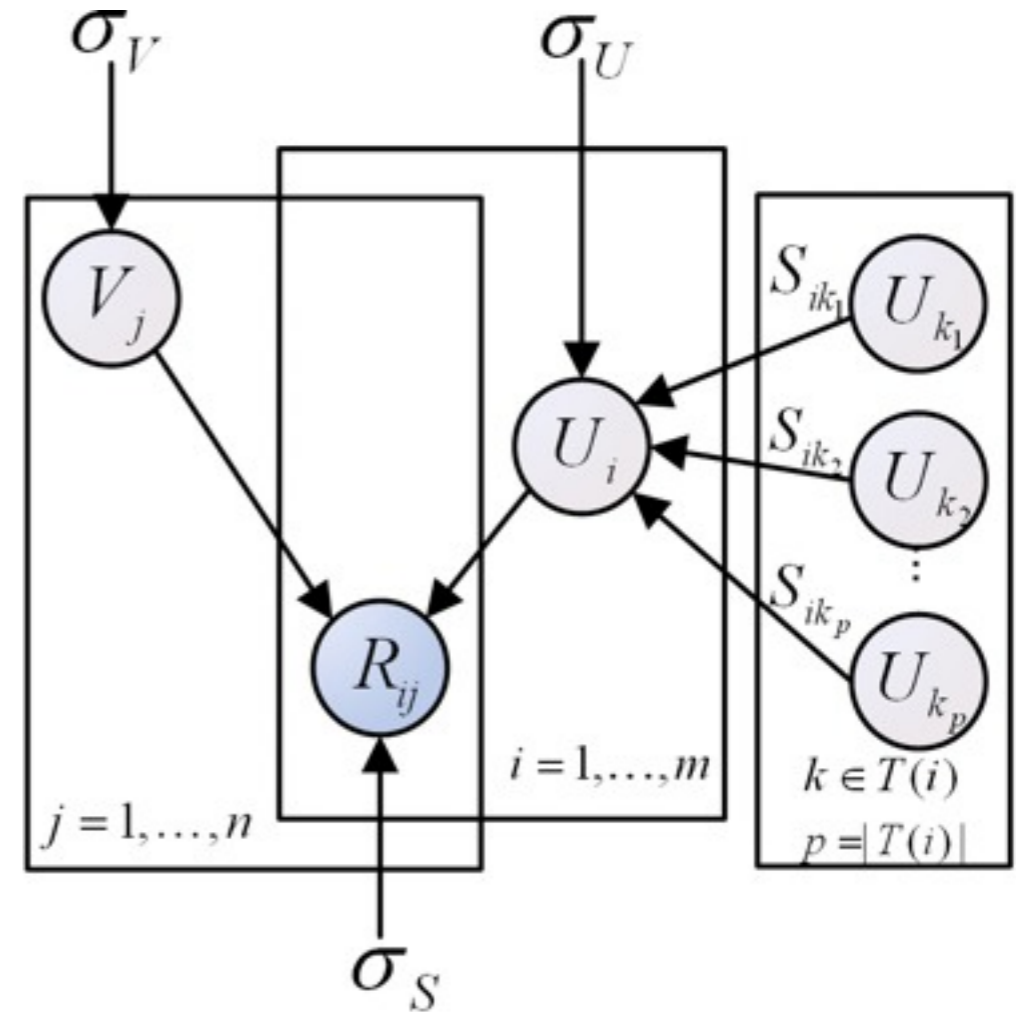
Recommendations by Trusted Friends

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

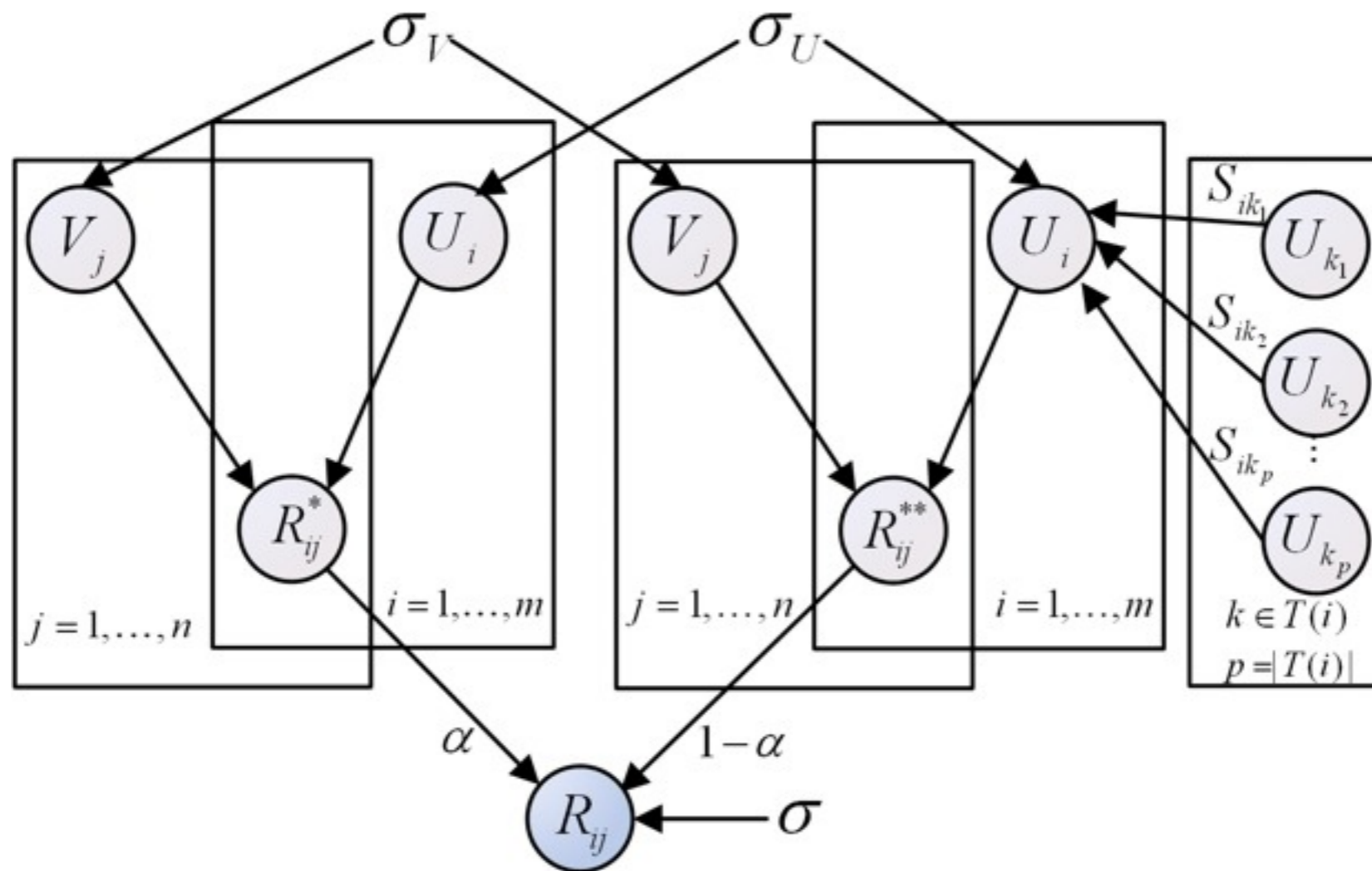
$$\hat{R}_{ik} = \sum_{j \in \mathcal{T}(i)} R_{jk} S_{ij}$$

$$p(R|S, U, V, \sigma_R^2) =$$

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



Recommendation with Social Trust Ensemble



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



Recommendation with Social Trust Ensemble

$$\begin{aligned}
 \mathcal{L}(R, S, U, V) &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (R_{ij} - g(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j))^2 \\
 &\quad + \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2,
 \end{aligned} \tag{15}$$

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

$$\begin{aligned}
 \frac{\partial \mathcal{L}}{\partial V_j} &= \sum_{i=1}^m I_{ij}^R g'(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) \\
 &\quad \times (g(\alpha U_i^T V_j + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) - R_{ij}) \\
 &\quad \times (\alpha U_i + (1 - \alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T) + \lambda_V V_j,
 \end{aligned}$$



Recommend with Social Distrust

[Ma et al., RecSys2009]



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



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



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

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

$$\begin{aligned} \min_{U, V} \mathcal{L}_{\mathcal{D}}(R, S^{\mathcal{D}}, U, V) &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (R_{ij} - g(U_i^T V_j))^2 \\ &+ \frac{\beta}{2} \sum_{i=1}^m \sum_{d \in \mathcal{D}^+(i)} (-S_{id}^{\mathcal{D}} \|U_i - U_d\|_F^2) \\ &+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2. \end{aligned}$$



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

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

$$\begin{aligned} \min_{U,V} \mathcal{L}_{\mathcal{T}}(R, S^T, U, V) &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^R (R_{ij} - g(U_i^T V_j))^2 \\ &+ \frac{\alpha}{2} \sum_{i=1}^m \sum_{t \in \mathcal{T}^+(i)} (S_{it}^T \|U_i - U_t\|_F^2) \\ &+ \frac{\lambda_U}{2} \|U\|_F^2 + \frac{\lambda_V}{2} \|V\|_F^2. \end{aligned}$$



Web Site Recommendation

[Ma et al., SIGIR 2011]



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



Traditional Search Paradigm

The image shows a screenshot of a Bing search results page for the query "sigir". The page layout includes a top navigation bar with links for "Web Images Videos Shopping News Maps More | MSN Hotmail", a user profile for "Irwin" with a Facebook icon, "Sign out", "Rewards", and location information "Walnut Creek, California" and "Preferences". The Bing logo is on the left, and the search bar contains "sigir". Below the search bar are tabs for "Web", "News", "Images", and "More".

RELATED SEARCHES

- Special Inspector General for Iraq Reconstruction
- SIGIR Reports
- SIGIR Poster
- SIGIR List
- SIGIR 2011
- SIGIR 10
- SIGIR 2010 Registration
- SIGIR 2009 Proceedings

SEARCH HISTORY

Search more to see your history

See all
Clear all · Turn off

NARROW BY DATE

- All results
- Past 24 hours
- Past week
- Past month

ALL RESULTS 1-10 of 255,000 results · [Advanced](#)

Welcome to SIGIR | Home
An Iraqi fisherman pushes his boat off-shore to depart on his daily fishing trip. View the Report.
www.sigir.mil

ACM SIGIR Special Interest Group on Information Retrieval Home Page
Welcome to the ACM SIGIR Web site. ACM SIGIR addresses issues ranging from theory to user demands in the application of computers to the acquisition, organization ...
www.sigir.org

home [ACM SIGIR 2010]
ACM-SIGIR 2010 was held at UniMail, Geneva, Switzerland between 19th and 23rd of July 2010. Thanks to all the participants!!! The story continues with ACM-SIGIR 2011.
www.sigir2010.org

Welcome to The 34th Annual ACM SIGIR Conference
Important Dates. 17 Jan 2011 : Abstracts for full research papers due; 24 Jan 2011 : Full research paper submissions due; 28 Jan 2011 : Workshop proposals due
sigir2011.org

About SIGIR
About SIGIR The Office of the Special Inspector General for Iraq Reconstruction (SIGIR) is the successor to the Coalition Provisional Authority Office of ...
www.sigir.mil/about/index.html

SIGIR 2009 Archive | SIGIR'09
The SIGIR 2009 conference ran July 19-23, 2009, in Boston, Massachusetts, at the Sheraton Boston Hotel and Northeastern University. The conference was chock full of ...
sigir2009.org

Bing Rewards

Earn Rewards with Bing
Join Bing Rewards for free and earn 250 credits.



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



“Search” to “Discovery”



[NEW Corp - Extended Service contracts, Extended warranties ...](#)
NEW delivers innovative extended service plans, also known as extended warranties, and customizable lifecycle solutions for the entire consumer ownership experience ...
<https://www.newcorp.com>

Careers Consumer
Contact Us Manufacturer
About Us Wireless
Service Plan FAQs Customer Care
Show more results from www.newcorp.com

[News Corporation](#)
07.15.2011. Les Hinton, Chief Executive Officer of Dow Jones & Company and Publisher ... View All News Corp. Press Releases >>
www.newscorp.com

[NEW Corp :: Careers](#)
Be Part of Something NEW. Founded in 1983, NEW has built a world-class organization dedicated to providing innovative and comprehensive customer care solutions and delivering ...
<https://www.newcorp.com/index.php/careers>



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

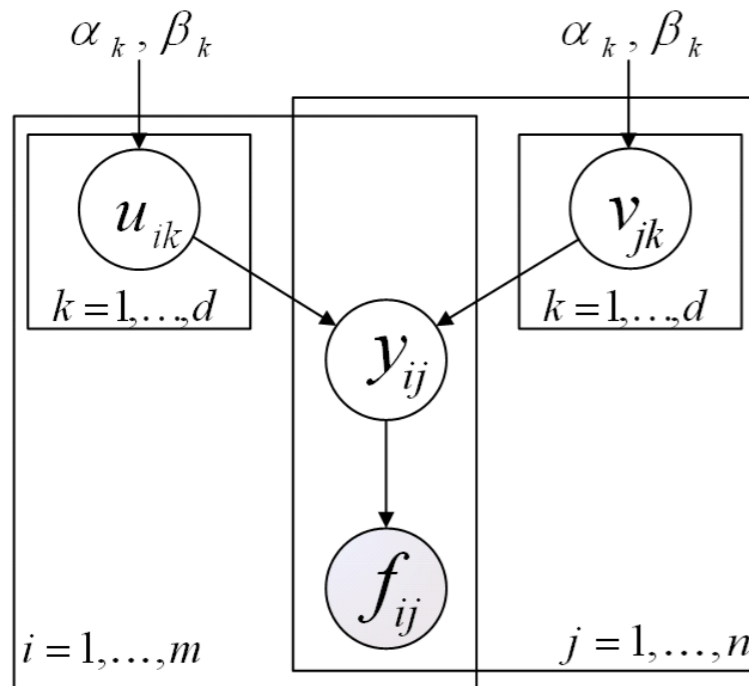
ID	Query	URL
358	facebook	http://www.facebook.com
358	rww	http://www.readwriteweb.com
3968	iphone4	http://www.apple.com
3968	ipad	http://www.apple.com
...

		Web sites					
		v_1	v_2	v_3	v_4	v_5	v_6
Web users	u_1		68	1		15	
	u_2	42			13		24
	u_3		72	12		11	2
	u_4	15			33		
	u_5		85	45			63

		Queries				
		z_1	z_2	z_3	z_4	z_5
Web users	u_1	12		5	6	
	u_2		23		5	1
	u_3		14		35	18
	u_4	25		11	4	
	u_5		12	5		24



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

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

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

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

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

$$\begin{aligned} \mathcal{L}(U, V; F) = & \sum_{i=1}^m \sum_{k=1}^d ((\alpha_k - 1) \ln(u_{ik}/\beta_k) - u_{ik}/\beta_k) \\ & + \sum_{j=1}^n \sum_{k=1}^d ((\alpha_k - 1) \ln(v_{jk}/\beta_k) - v_{jk}/\beta_k) \\ & + \sum_{i=1}^m \sum_{j=1}^n (f_{ij} \ln y_{ij} - y_{ij}) + \text{const.} \end{aligned}$$



Probabilistic Factor Model

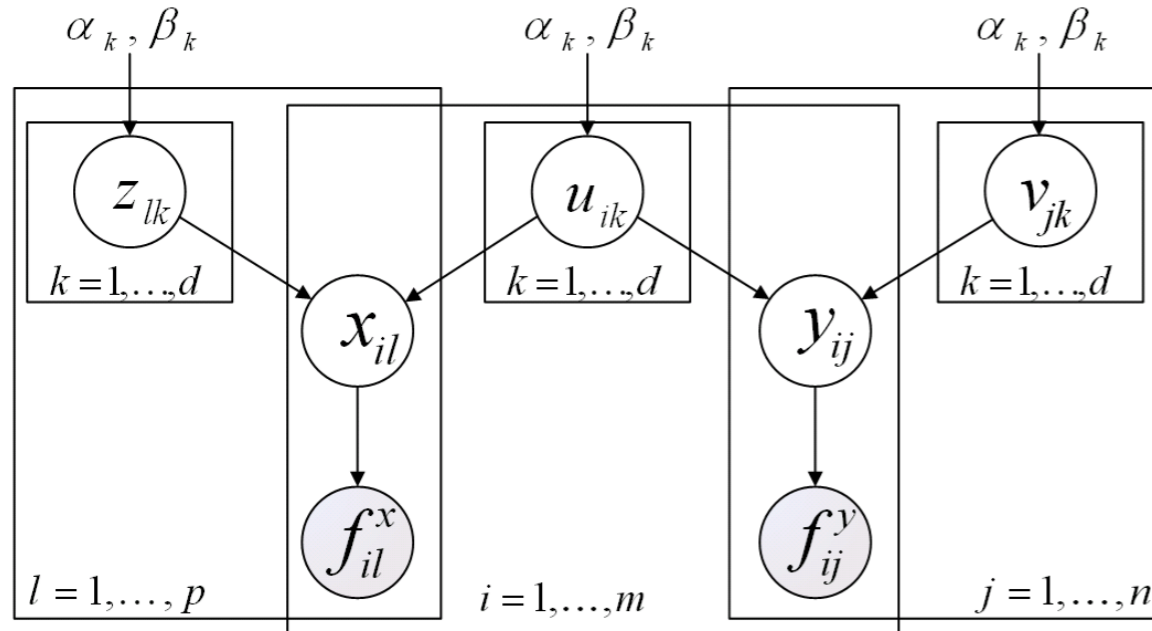
$$\begin{aligned}\mathcal{L}(U, V; F) &= \sum_{i=1}^m \sum_{k=1}^d ((\alpha_k - 1) \ln(u_{ik}/\beta_k) - u_{ik}/\beta_k) \\ &\quad + \sum_{j=1}^n \sum_{k=1}^d ((\alpha_k - 1) \ln(v_{jk}/\beta_k) - v_{jk}/\beta_k) \\ &\quad + \sum_{i=1}^m \sum_{j=1}^n (f_{ij} \ln y_{ij} - y_{ij}) + \text{const.}\end{aligned}$$

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

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



Collective Probabilistic Factor Model



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

$$\begin{aligned}
 &= \sum_{i=1}^m \sum_{l=1}^p (f_{il}^x \ln x_{il} - x_{il}) + \sum_{i=1}^m \sum_{j=1}^n (f_{ij}^y \ln y_{ij} - y_{ij}) \\
 &+ \sum_{i=1}^m \sum_{k=1}^d ((\alpha_k - 1) \ln(u_{ik}/\beta_k) - u_{ik}/\beta_k) \\
 &+ \sum_{j=1}^n \sum_{k=1}^d ((\alpha_k - 1) \ln(v_{jk}/\beta_k) - v_{jk}/\beta_k) \\
 &+ \sum_{l=1}^p \sum_{k=1}^d ((\alpha_k - 1) \ln(z_{lk}/\beta_k) - z_{lk}/\beta_k) + \text{const.}
 \end{aligned}$$

$$u_{ik} \leftarrow u_{ik} \frac{\sum_{j=1}^n (f_{ij}^y v_{jk}/y_{ij}) + \sum_{l=1}^p (f_{il}^x z_{lk}/x_{il}) + (\alpha_k - 1)/u_{ik}}{\sum_{j=1}^n v_{jk} + \sum_{l=1}^p z_{lk} + 1/\beta_k}$$

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

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

$$u_{ik} \leftarrow u_{ik} \frac{\theta \sum_{j=1}^n (f_{ij}^y v_{jk}/y_{ij}) + (1-\theta) \sum_{l=1}^p (f_{il}^x z_{lk}/x_{il}) + (\alpha_k - 1)/u_{ik}}{\theta \sum_{j=1}^n v_{jk} + (1-\theta) \sum_{l=1}^p z_{lk} + 1/\beta_k}$$



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

Statistics	User-Site Frequency	User-Query Frequency
Min. Num.	4	10
Max. Num.	9,969	4,693
Avg. Num.	20.33	23.05



Performance Comparison

Table 3: Performance Comparison (Dimensionality = 10)

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

Table 4: Performance Comparison (Dimensionality = 20)

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



Impact of Parameters

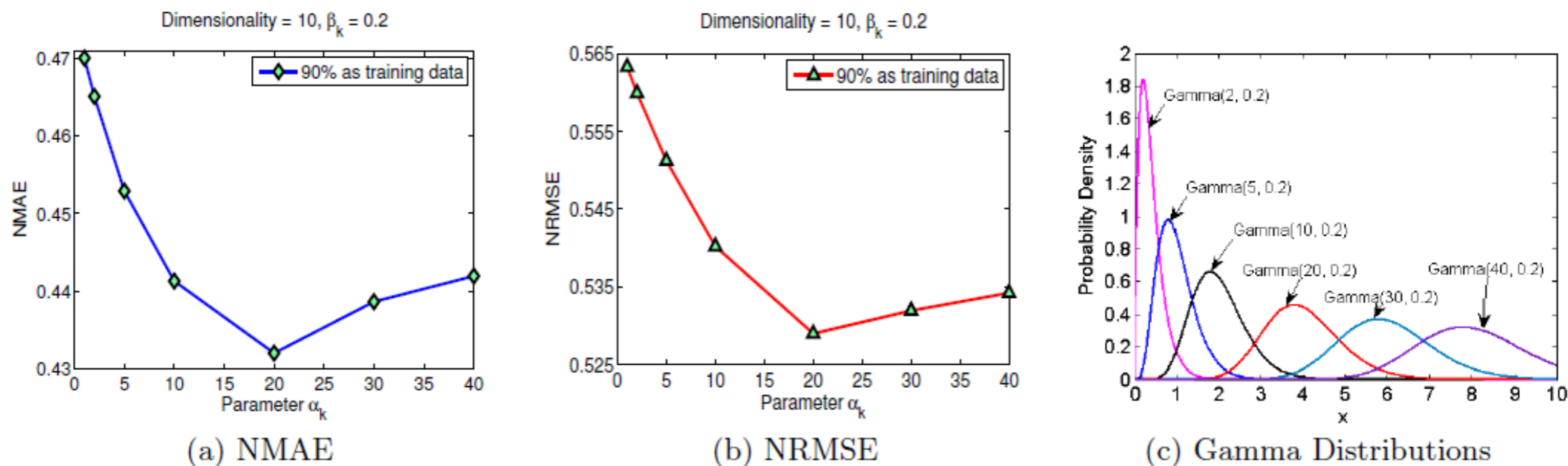


Figure 6: Impact of Parameter α_k in PFM

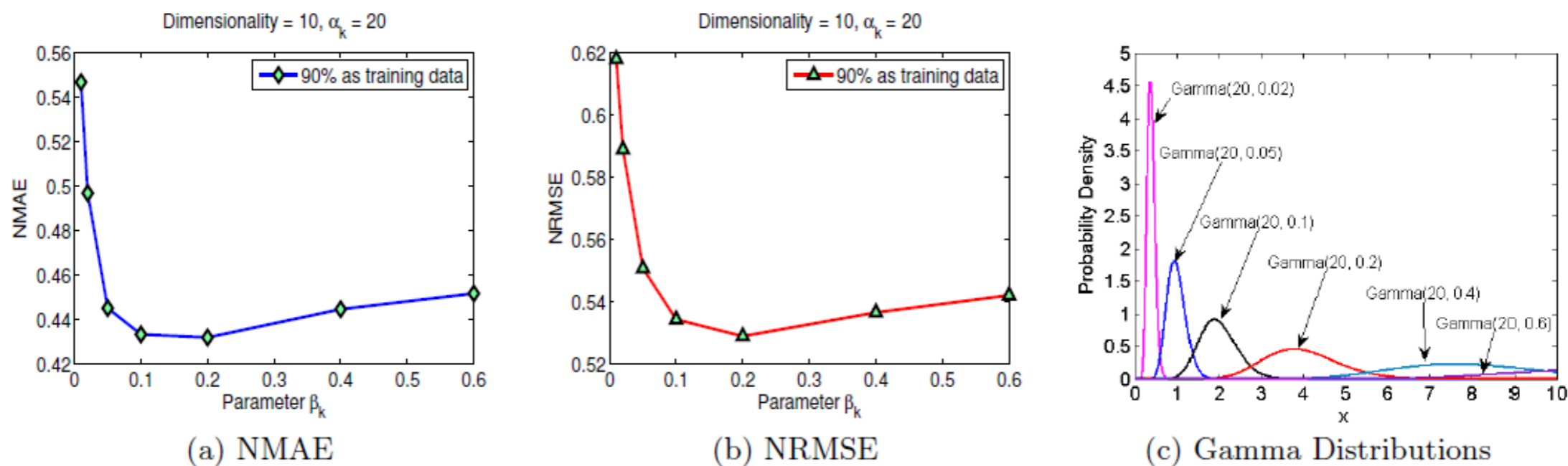
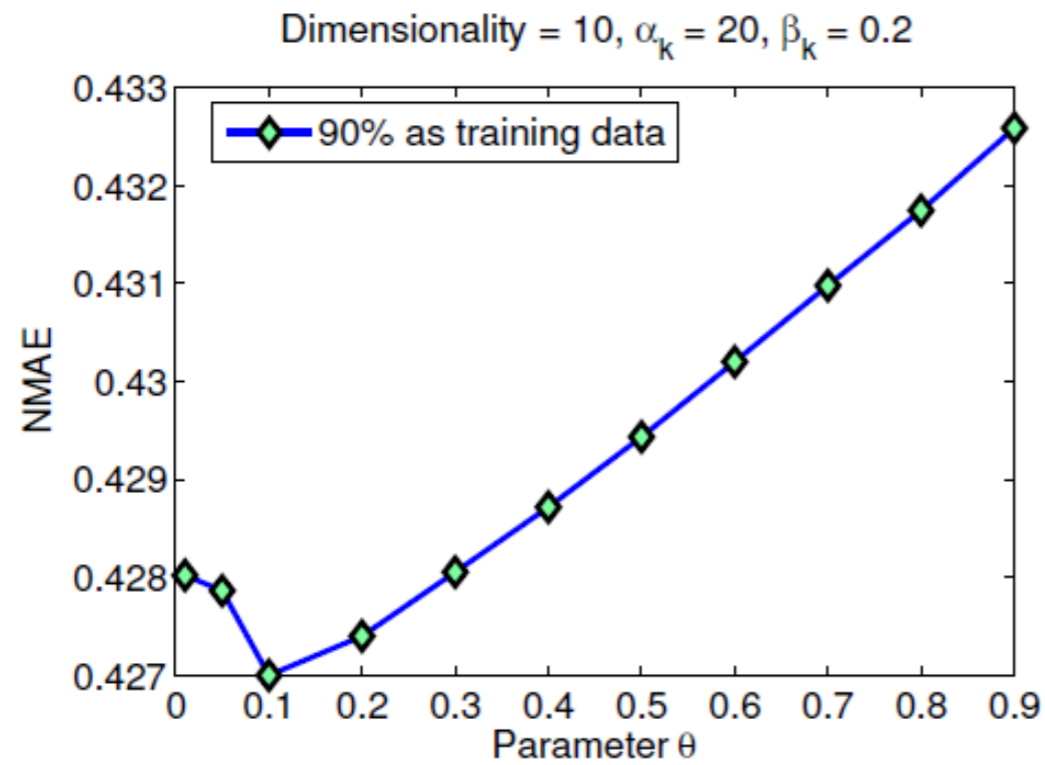


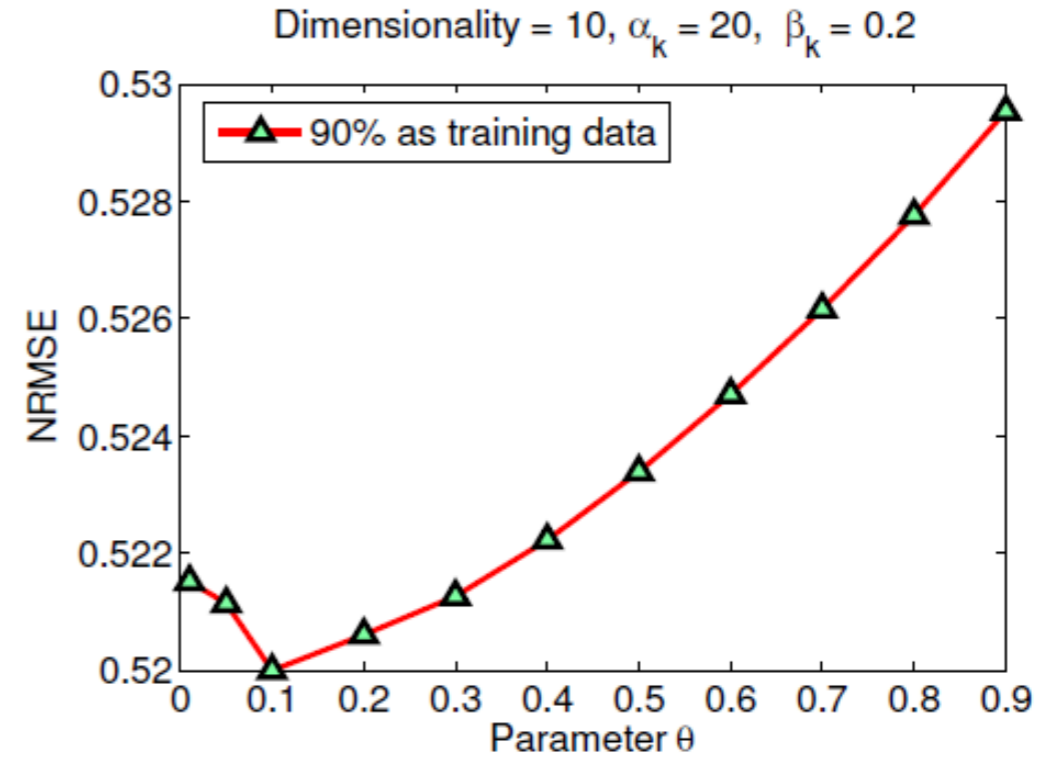
Figure 7: Impact of Parameter β_k in PFM



Impact of Parameters



(a) NMAE



(b) NRMSE

Figure 8: Impact of Parameter θ in CPM



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.)



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



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](#))



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



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](#))



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



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](#))



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



Irwin King
Ricardo Baeza-Yates (Eds.)

King · Baeza-Yates (Eds.)



Weaving Services and People
on the World Wide Web

Weaving Services and People on the World Wide Web

King · Baeza-Yates (Eds.)

Weaving Services and People on the World Wide Web

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.

ISBN 978-3-642-00569-5



9 783642 005695

springer.com



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



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



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

