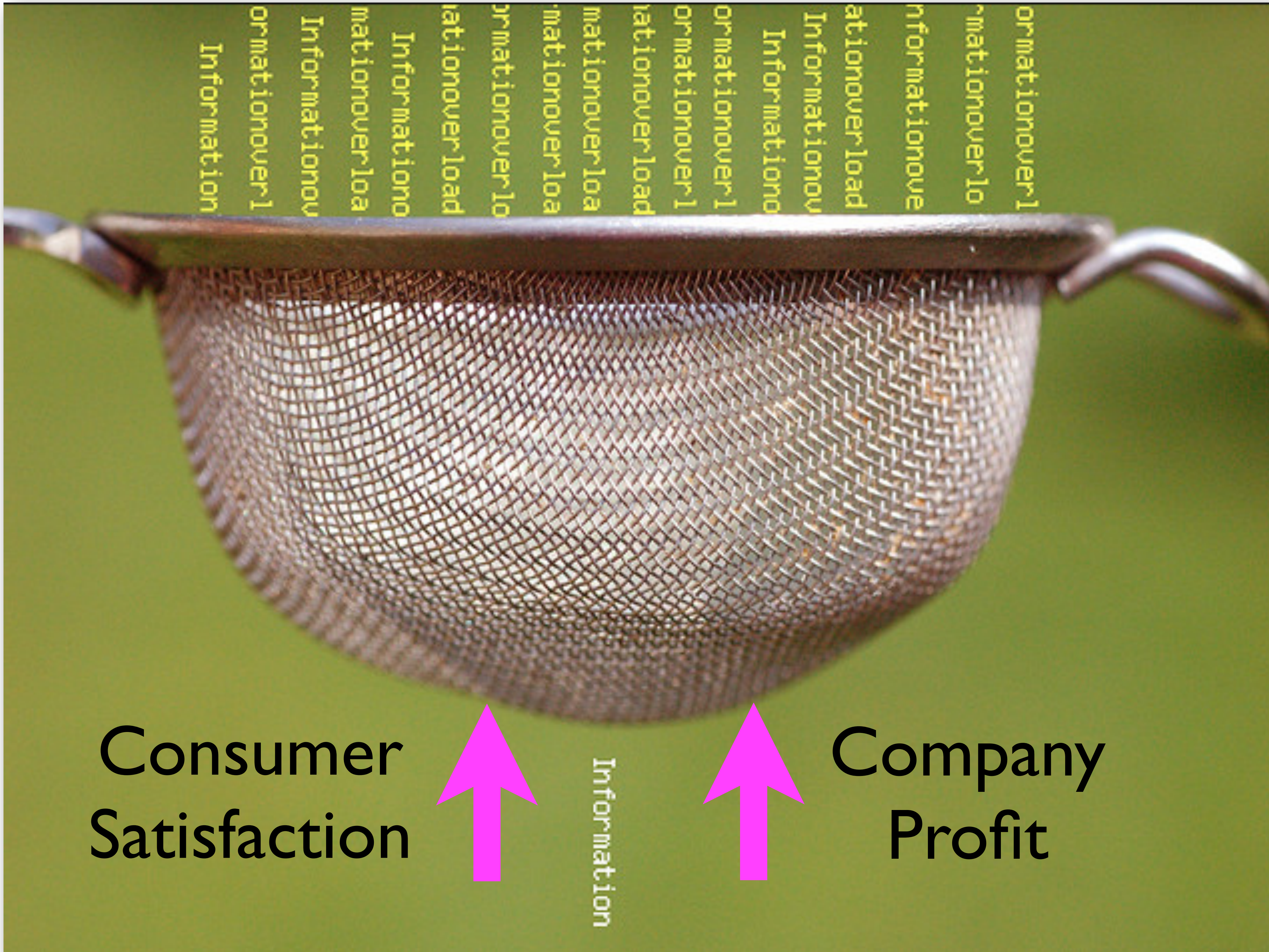
Recent Developments in Social and Location Recommendations

Irwin King
Joint work with Hao Ma and Cheng Chen

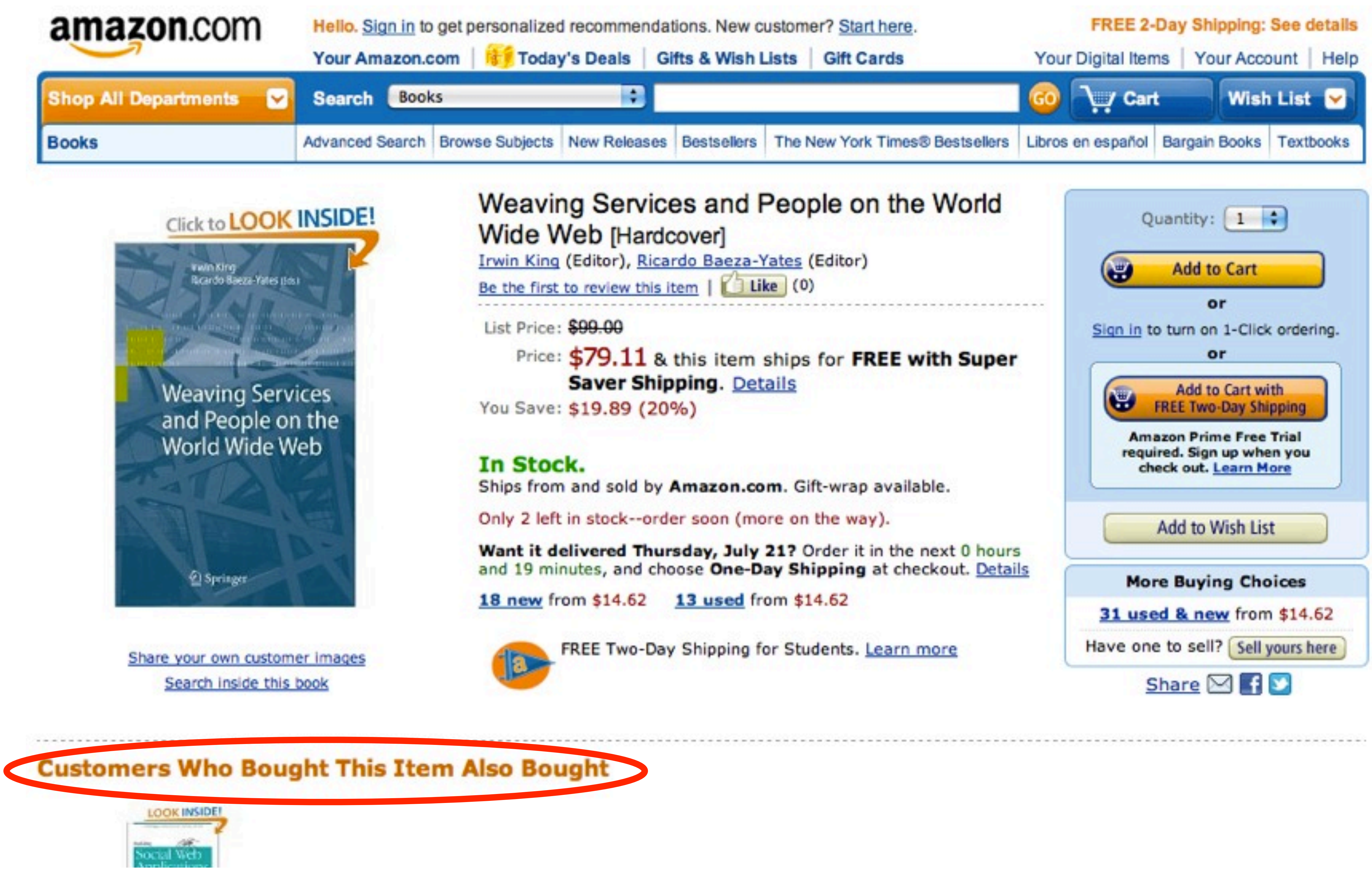
Department of Computer Science and Engineering
The Chinese University of Hong Kong

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

Information and More Information!

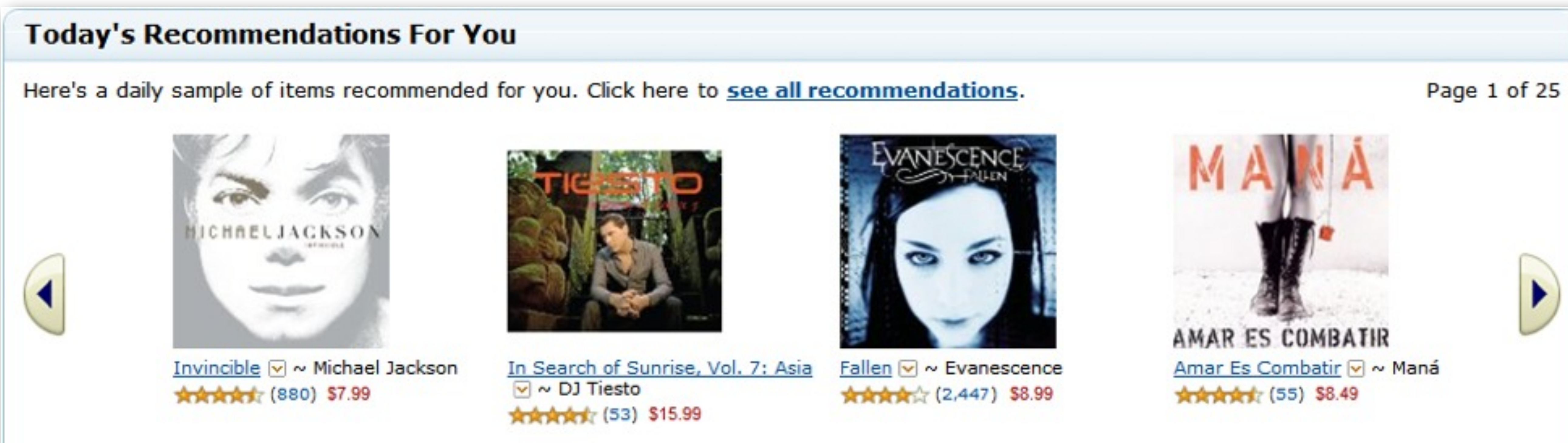

Consumer Satisfaction
Company Profit
Information

Real Life Examples



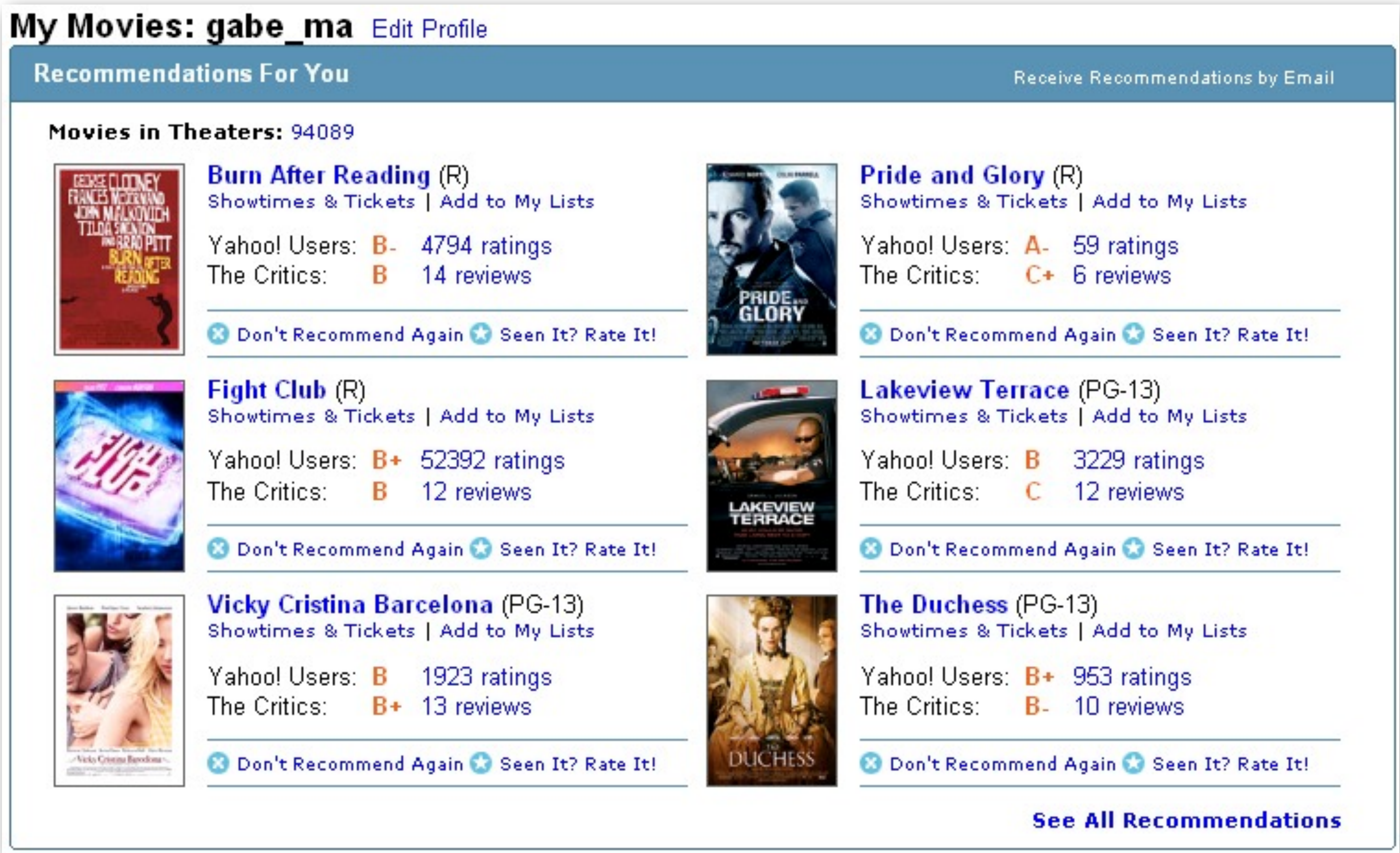
Real Life Examples

Real Life Examples



Real Life Examples

YAHOO! MOVIES



Real Life Examples

iTunes 8

On The Menu

- Introduction
- Social Recommendation Models
    - Social graph
    - Social ensemble
    - Social distrust
    - Website recommendation
- Multi-centered Gaussian Location Recommendation Model
- Conclusion

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 other similar users prefer!

Framework

Items

Users

	i_1	i_2				i_j							i_m
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 should be recommended?

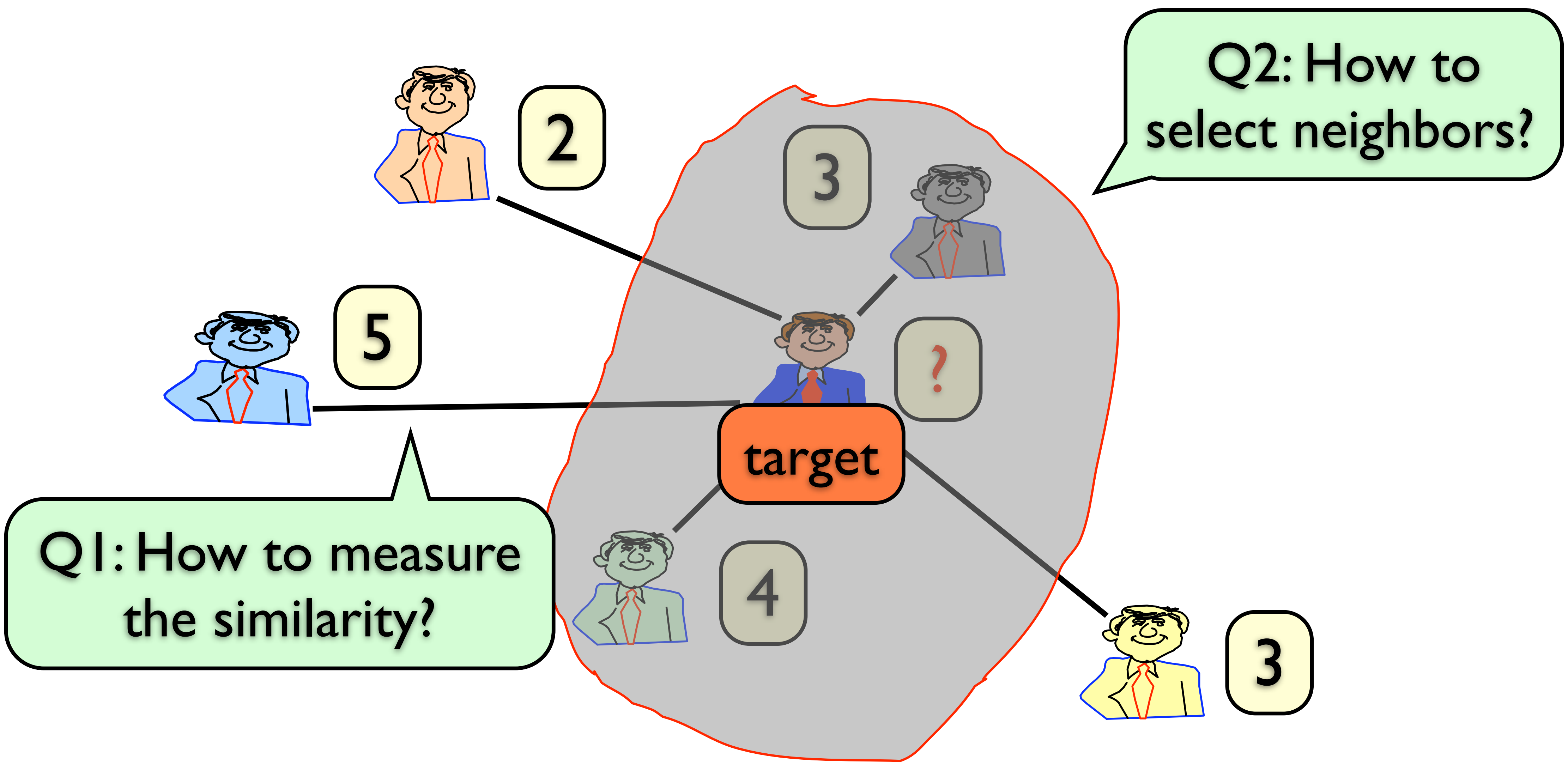
Social Recommender Systems

- Introduction
- Collaborative Filtering
- Trust-aware Recommender Systems
- Social-based Recommender Systems

Collaborative Filtering

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

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]

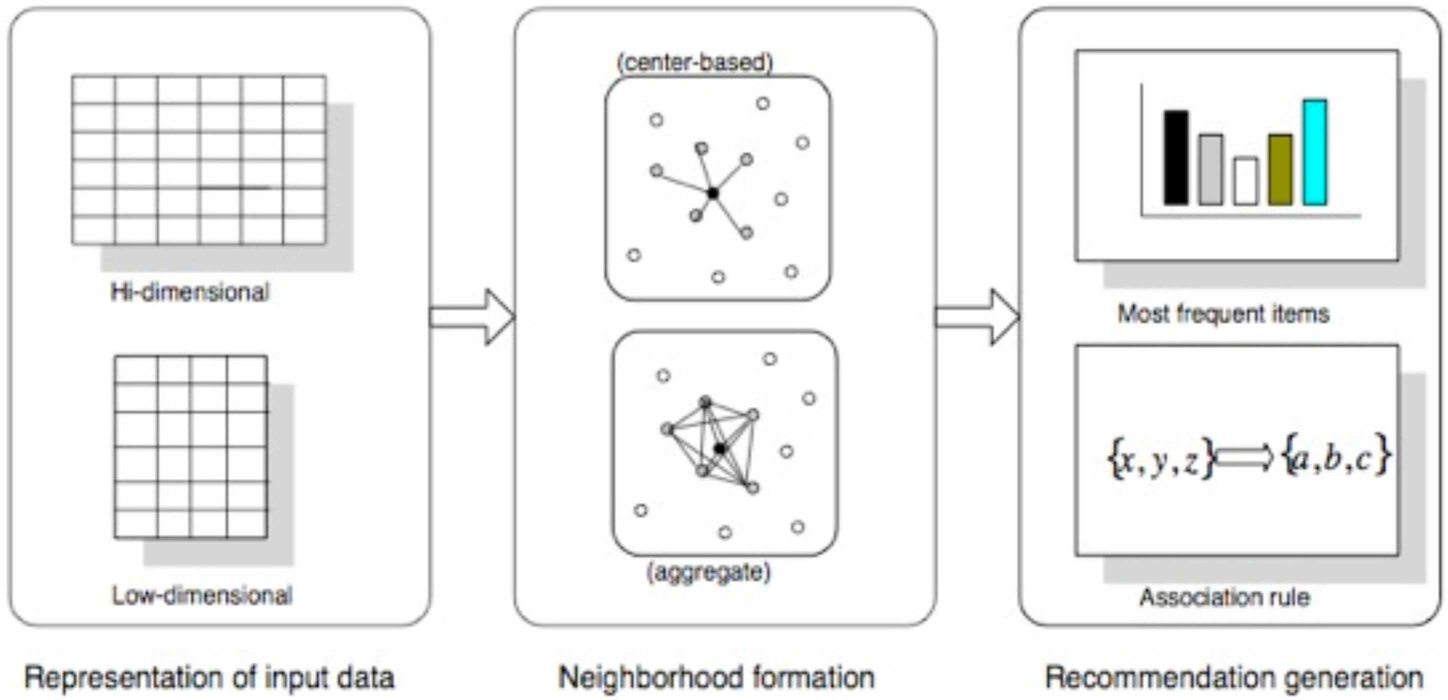


Figure 1: Three main parts of a Recommender System.

- Identify highly similar users to the active one
  - All with a measure greater than a threshold
  - Best K ones
- Prediction $r_{aj} = \bar{r}_a + \frac{\sum_i w(a,i)(r_{ij} - \bar{r}_i)}{\sum_i w(a,i)}$

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.

Item-Item Similarity

- Search for similarities among items
- Item-Item similarity is more stable than user-user similarity

Correlation-based Methods

[Sarwar, 2001]

- Same as in user-user similarity but on item vectors
- Pearson correlation coefficient
  - Look for users who rated both items

$$s_{ij} = \frac{\sum_u (r_{uj} - \bar{r}_j)(r_{ui} - \bar{r}_i)}{\sqrt{\sum_u (r_{uj} - \bar{r}_j)^2 \sum_u (r_{ui} - \bar{r}_i)^2}}$$

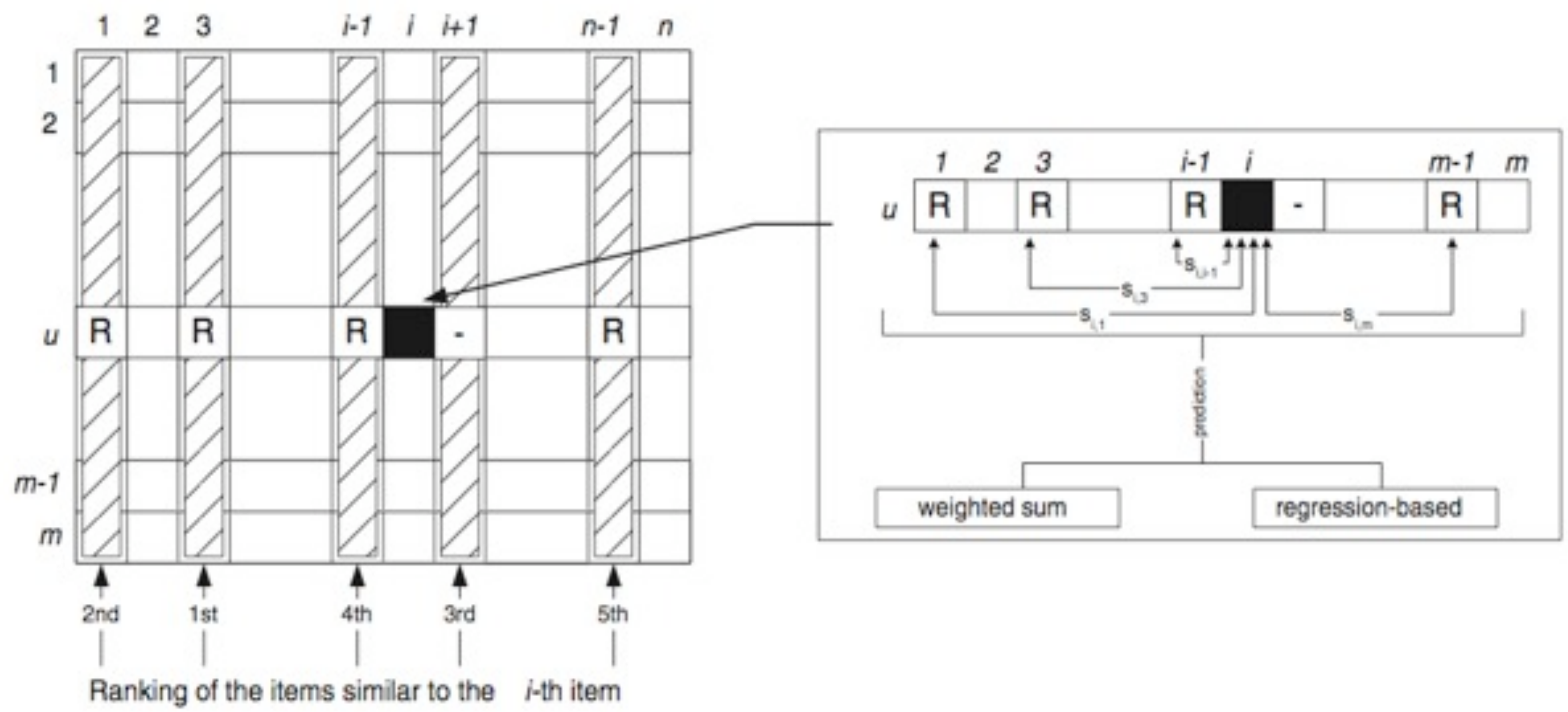
  - u: users rated both items

	i_1	i_2				i_i		i_j					i_m
u_1													
u_2	1	3		4		2		5			3	4	
u_i		3		4			3	4		3	4		4
u_n	1			3	5	2		4	1			3	

Correlation-based Method

[Sarwar, 2001]

- Calculate item similarity, then determine its k-most similar items



- Predict rating for a given user-item pair as a weighted sum over similar items that he rated

$$r_{ai} = \frac{\sum_j s_{ij} r_{aj}}{\sum_j s_{ij}}$$

					3		
u_a	2		3		?		4
					i		

Traditional Methods

- Memory-based Methods (Neighborhood-based Method)
  - Pearson Correlation Coefficient
  - User-based, Item-based
  - Etc.
- Model-based Method
  - Matrix Factorizations
  - Etc.

	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-based Method

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	

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

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

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 x 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

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

Regularization terms

where $\lambda_1, \lambda_2 > 0$.

- The problem can be solved by simple gradient descent algorithm.

Regularized Matrix Factorization

- Algorithm for RMF
  - Not convex & local optimal
  - Gradient-decent algorithm
    - Gradient computation with randomly initialized *U* and *V*

$$\frac{\partial L}{\partial u_{il}} = \lambda u_{il} - \sum_{j|(i,j)\in S} (y_{ij} - \widehat{y_{ij}}) v_{jl}$$

$$\frac{\partial L}{\partial v_{il}} = \lambda v_{il} - \sum_{j|(i,j)\in S} (y_{ij} - \widehat{y_{ij}}) u_{jl}$$

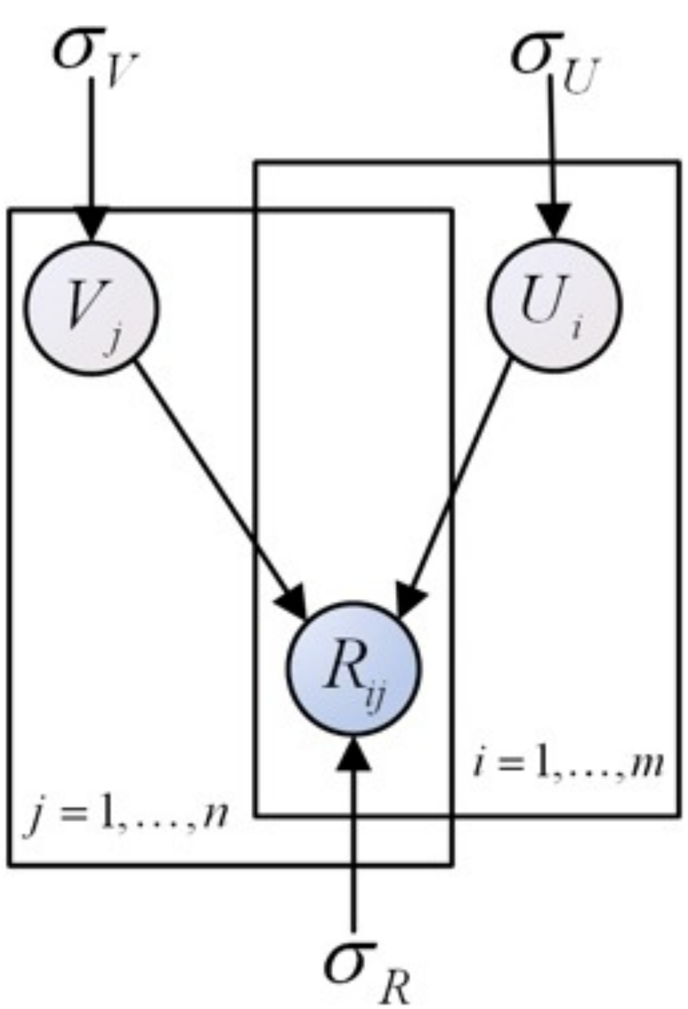
    - Update *U* and *V* alternatively

$$u_{il}^{(t+1)} = u_{il}^{(t)} - \tau \frac{\partial L}{\partial u_{il}^{(t)}}$$

$$v_{jl}^{(t+1)} = v_{jl}^{(t)} - \tau \frac{\partial L}{\partial v_{jl}^{(t)}}$$

Probabilistic Matrix Factorization

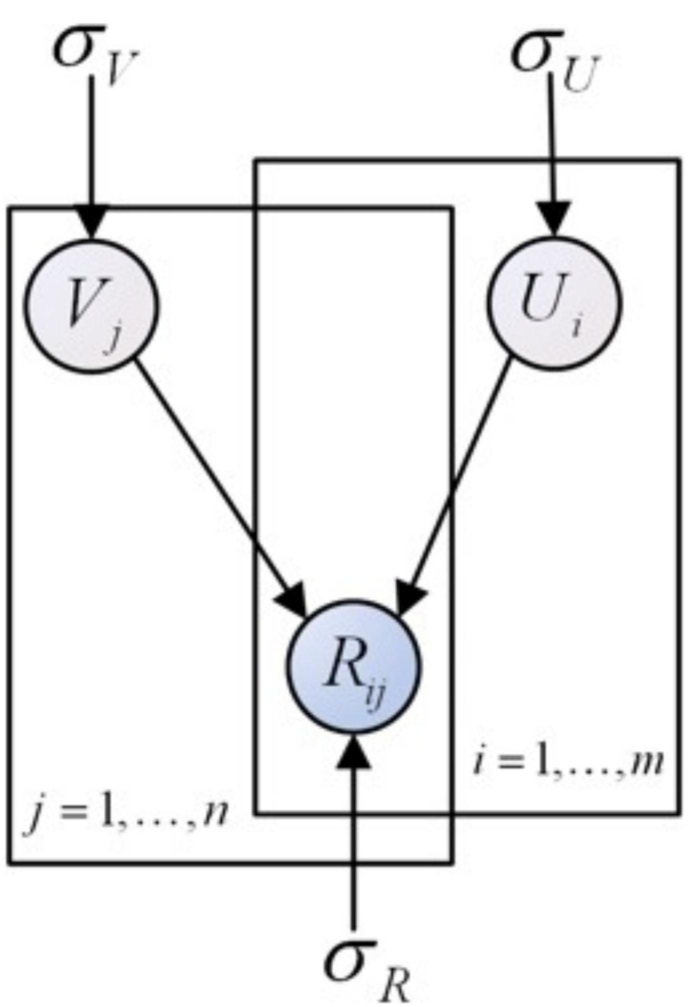
- PMF
  - Define a conditional distribution over the observed ratings as:



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

Probabilistic Matrix Factorization

- PMF
  - Assume zero-mean spherical Gaussian priors on user and item feature:

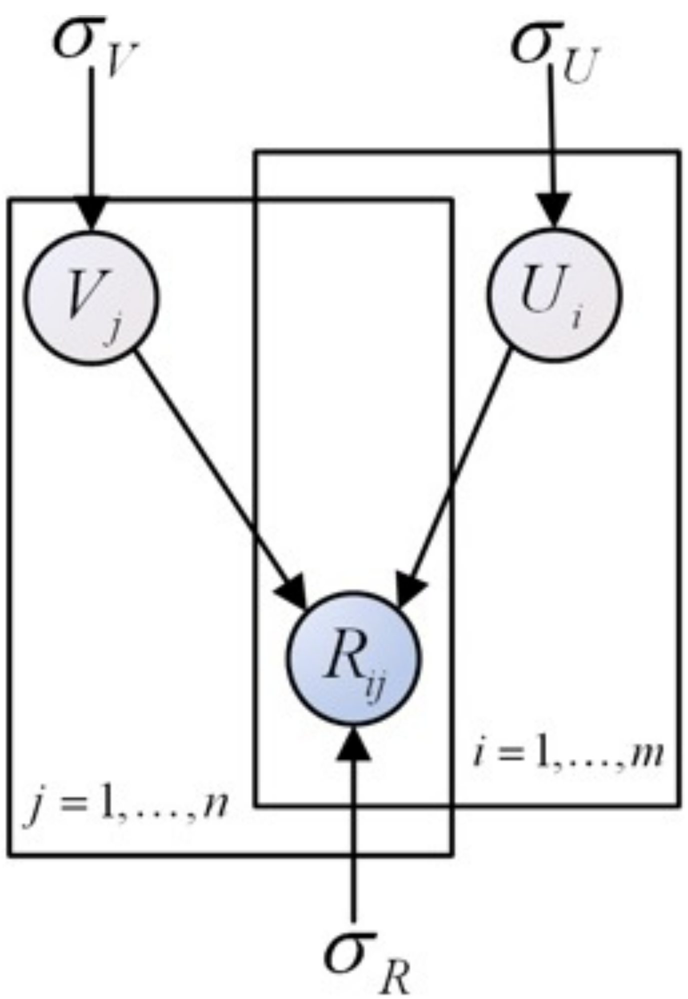


$$p(U|\sigma_U^2)=\prod_{i=1}^{m}\mathcal{N}(U_i|0,\sigma_U^2\mathbf{I})$$

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

Probabilistic Matrix Factorization

- PMF
  - Bayesian inference



$$
\begin{aligned}
p(U,V|R,\sigma_R^2,\sigma_U^2,\sigma_V^2) &\propto p(R|U,V,\sigma_R^2)p(U|\sigma_U^2)p(V|\sigma_V^2) \\
&= \prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathcal{N}\left(R_{ij}|g(U_i^T V_j),\sigma_R^2\right)\right]^{I_{ij}^R} \\
&\times \prod_{i=1}^{m}\mathcal{N}(U_i|0,\sigma_U^2\mathbf{I}) \times \prod_{j=1}^{n}\mathcal{N}(V_j|0,\sigma_V^2\mathbf{I}).
\end{aligned}
$$

RMF and PMF

- PMF is the probabilistic interpretation of RMF
- PMF and RMF have the same optimization objective function

Non-negative Matrix Factorization

- NMF
  - Non-negative constraints on all entries of matrices U and V

$$\begin{pmatrix} -10 & \cdots & u_{1d} \\ \vdots & \ddots & \vdots \\ u_{m1} & \cdots & u_{md} \end{pmatrix} \times \begin{pmatrix} -10 & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{d1} & \cdots & v_{dn} \end{pmatrix} \Rightarrow$$

Non-negativity Constraint

(-10)*(-10)=100

Non-negative Matrix Factorization

- NMF
  - Given an observed matrix Y, to find two non-negative matrices U and V
  - Two types of loss functions
    - Squared error function

$$\sum_{ij}(R_{ij} - U_i^T V_j)^2$$

    - Divergence

$$D(R||U^T V) = \sum_{ij}(R_{ij} \log \frac{R_{ij}}{U_i^T V_j} - R_{ij} + U_i^T V_j)$$

- Solving by multiplicative updating rules

Social Recommendation Using Probabilistic Matrix Factorization

[Ma et al., CIKM2008]

Challenges

- Traditional recommender systems ignore the social connections between users

Recommendations from friends

Motivations

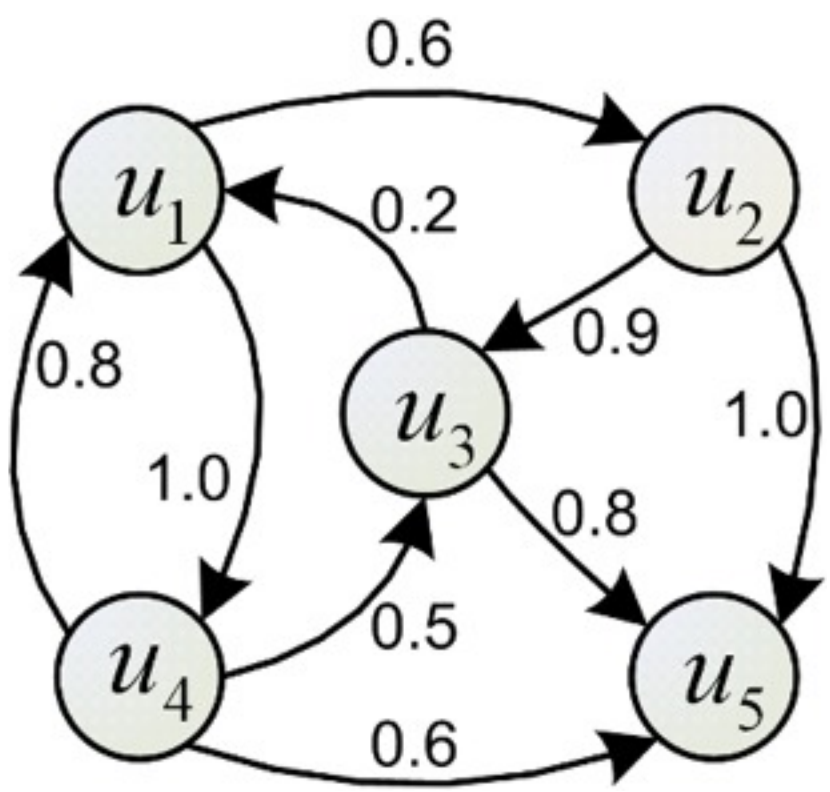
- "Yes, there is a correlation - from social networks to personal behavior on the web"

  Parag Singla and Matthew Richardson (WWW'08)

  - Analyze the who talks to whom social network over 10 million people with their related search results
  - People who chat with each other are more likely to share the same or similar interests

- To improve the recommendation accuracy and solve the data sparsity problem, users' social network should be taken into consideration

Problem Definition



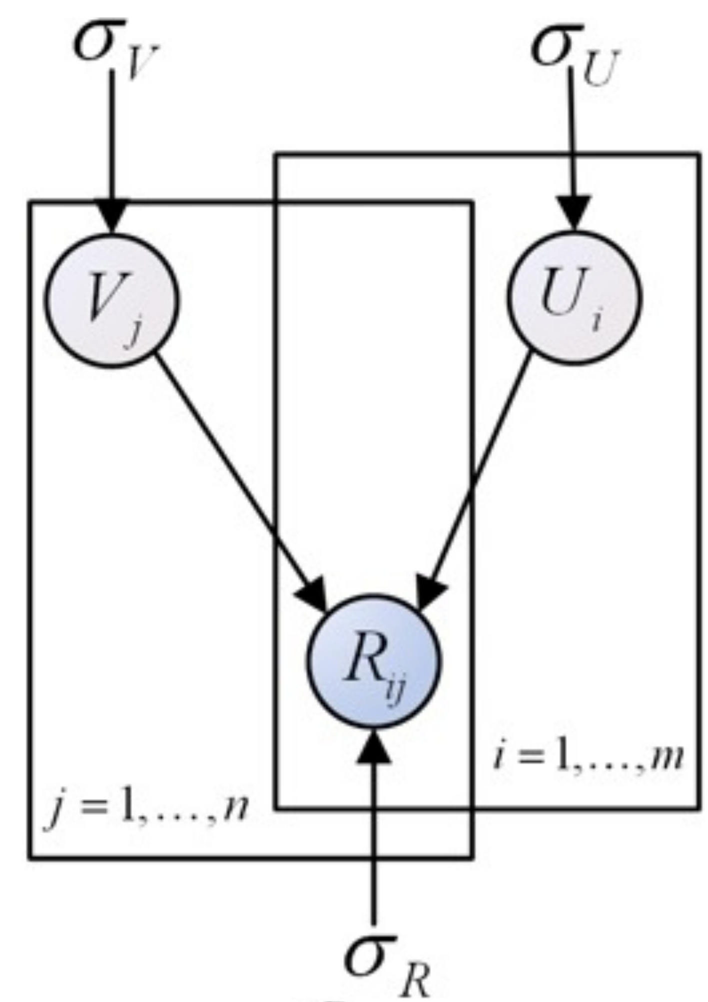
Social Trust Graph

	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



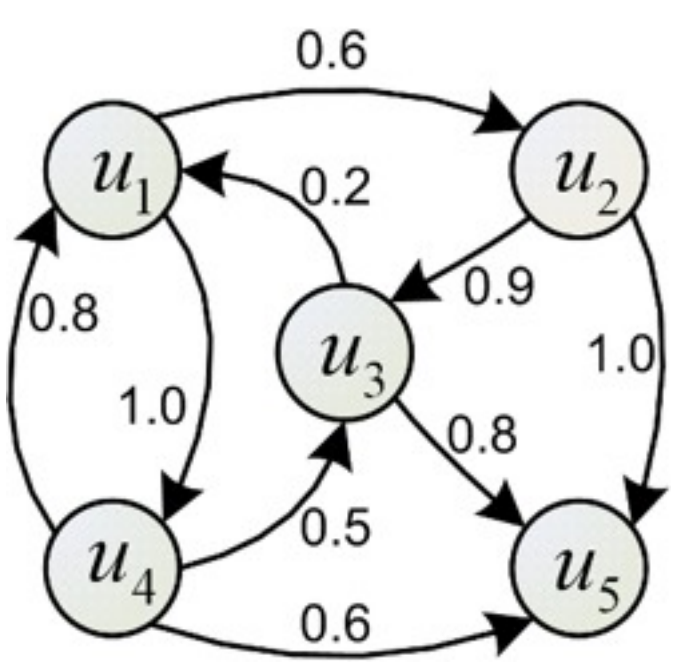
$$p(R|U,V,\sigma_R^2) = \prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathcal{N}\left(R_{ij}|g(U_i^T V_j),\sigma_R^2\right)\right]^{I_{ij}^R}$$

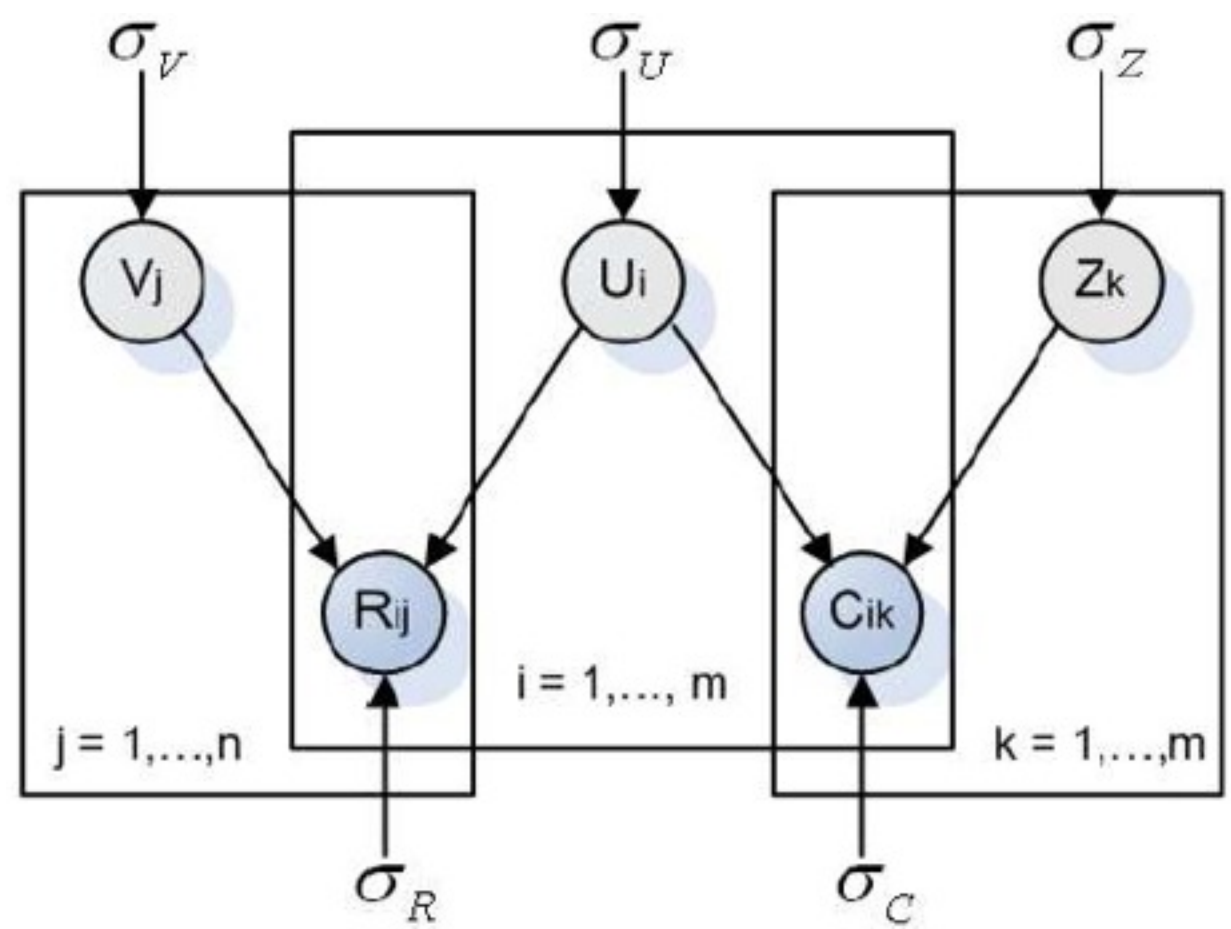
$$p(U|\sigma_U^2)=\prod_{i=1}^{m}\mathcal{N}(U_i|0,\sigma_U^2\mathbf{I}) \qquad p(V|\sigma_V^2)=\prod_{j=1}^{n}\mathcal{N}(V_j|0,\sigma_V^2\mathbf{I})$$

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

SoRec

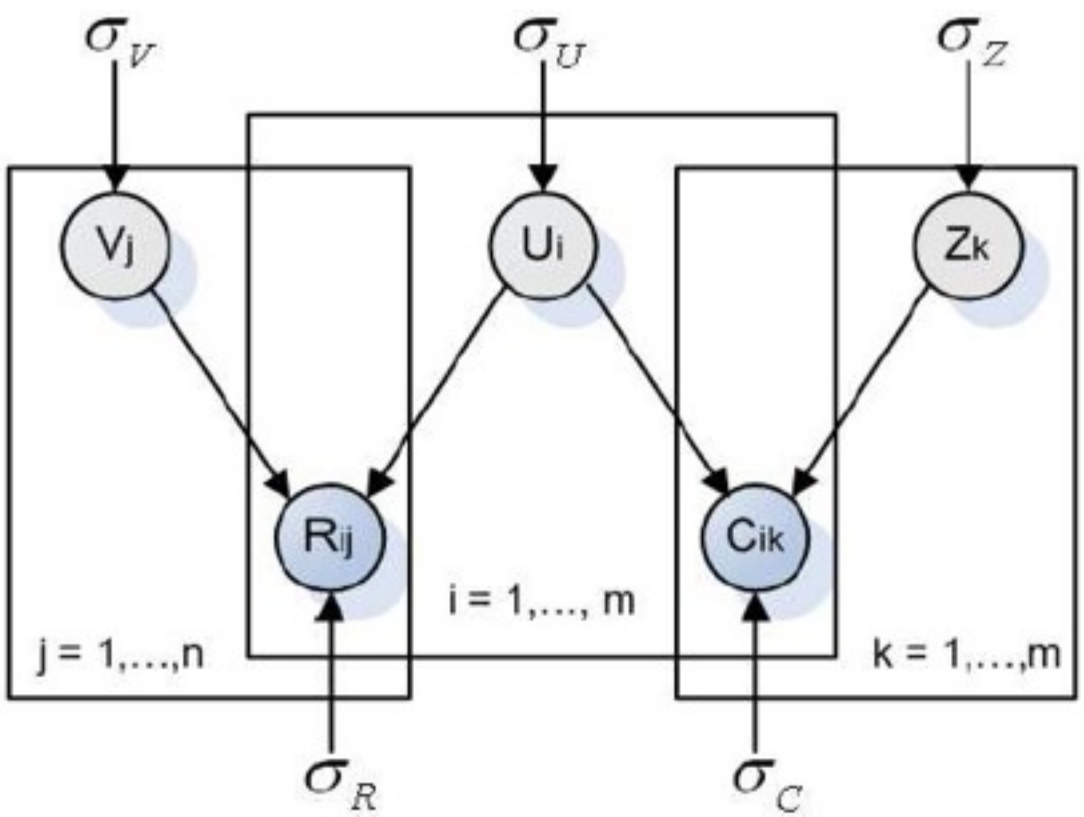
	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3





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

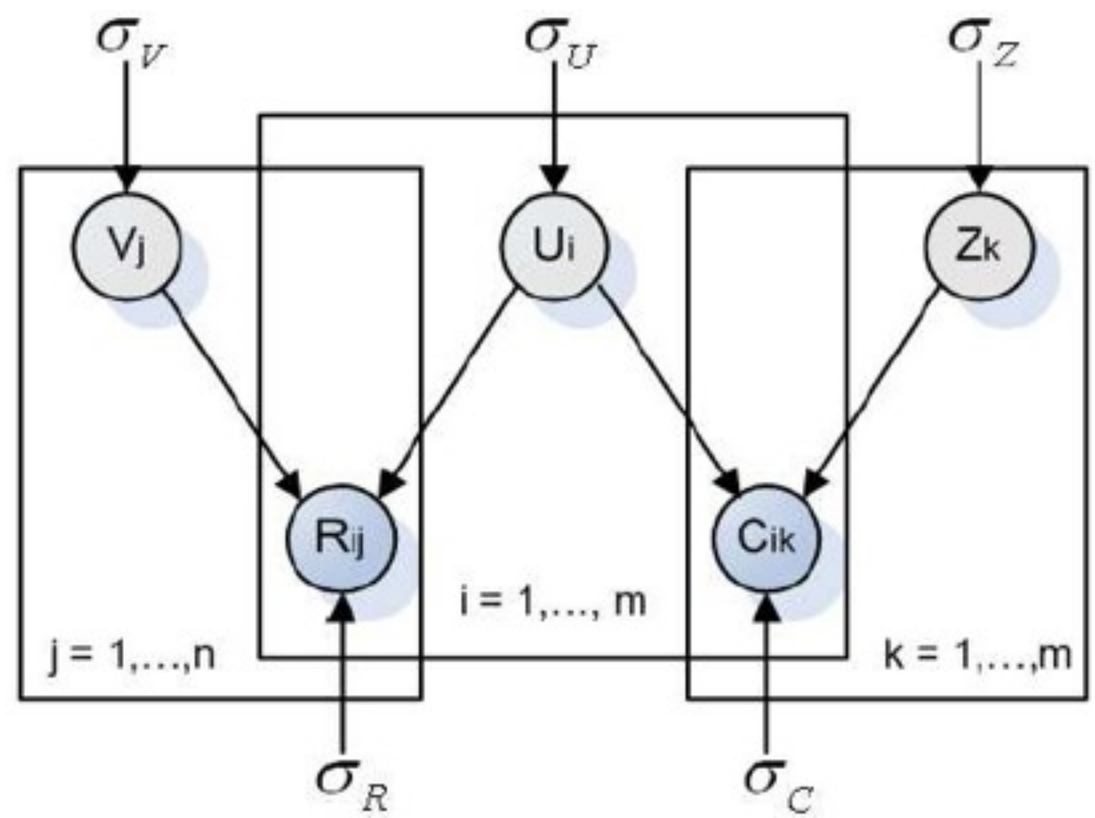
$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial U_i} &= \sum_{j=1}^{n} I_{ij}^{R} g'(U_i^T V_j)(g(U_i^T V_j) - r_{ij})V_j \\
&+ \lambda_C \sum_{j=1}^{m} I_{ik}^{C} g'(U_i^T Z_k)(g(U_i^T Z_k) - c_{ik}^{*})Z_k + \lambda_U U_i, \\
\frac{\partial \mathcal{L}}{\partial V_j} &= \sum_{i=1}^{m} I_{ij}^{R}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

Complexity Analysis

- For the Objective Function $O(\rho_R l + \rho_C l)$
- For $\frac{\partial \mathcal{L}}{\partial U}$ the complexity is $O(\rho_R l + \rho_C l)$
- For $\frac{\partial \mathcal{L}}{\partial V}$ the complexity is $O(\rho_R l)$
- For $\frac{\partial \mathcal{L}}{\partial Z}$ the complexity is $O(\rho_C l)$
- In general, the complexity of our method is linear with the observations in these two matrices

Disadvantages of SoRec

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



SoRec

Learning to Recommend with Social Trust Ensemble

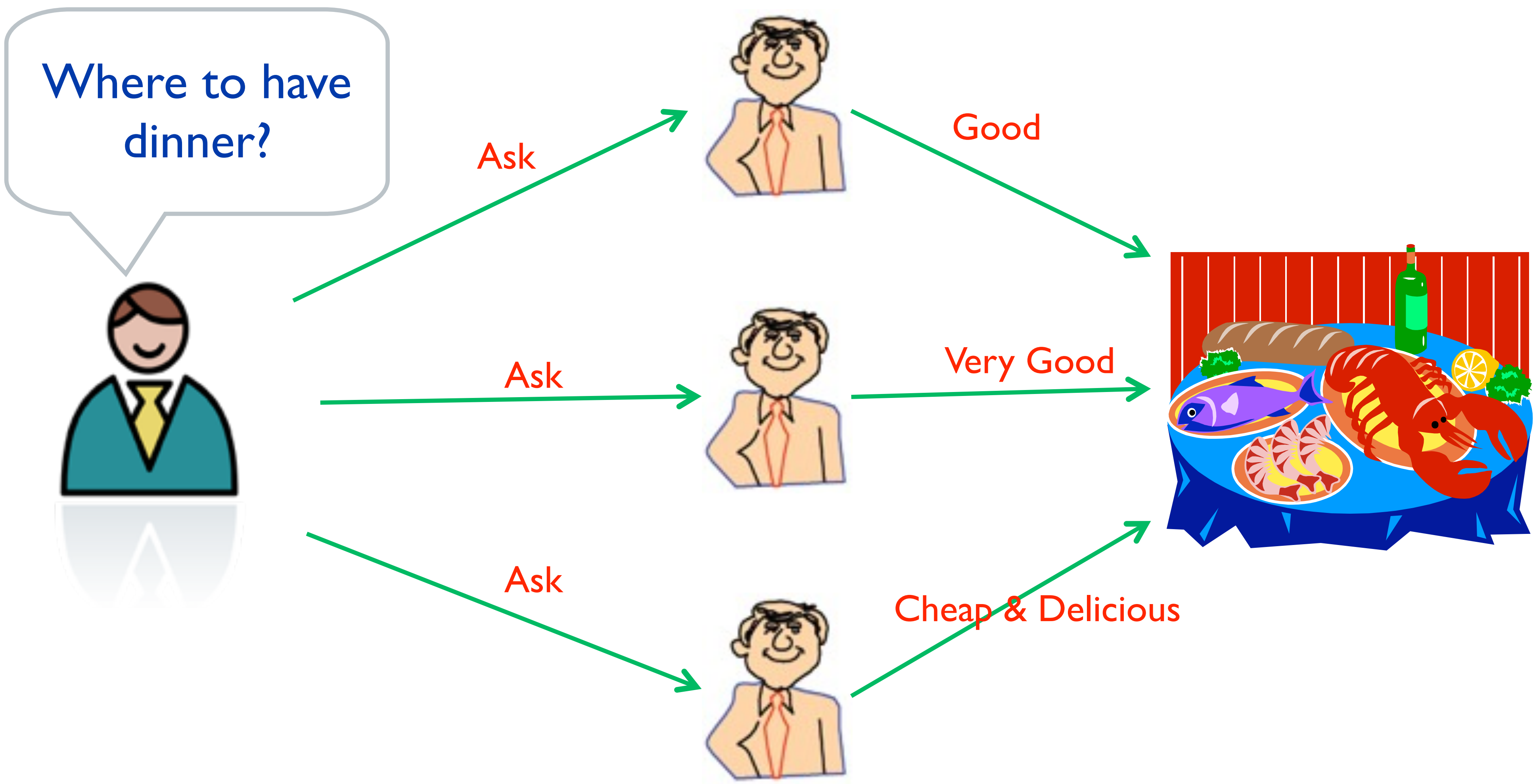
[Ma et al., SIGIR2009]

1st Motivation

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

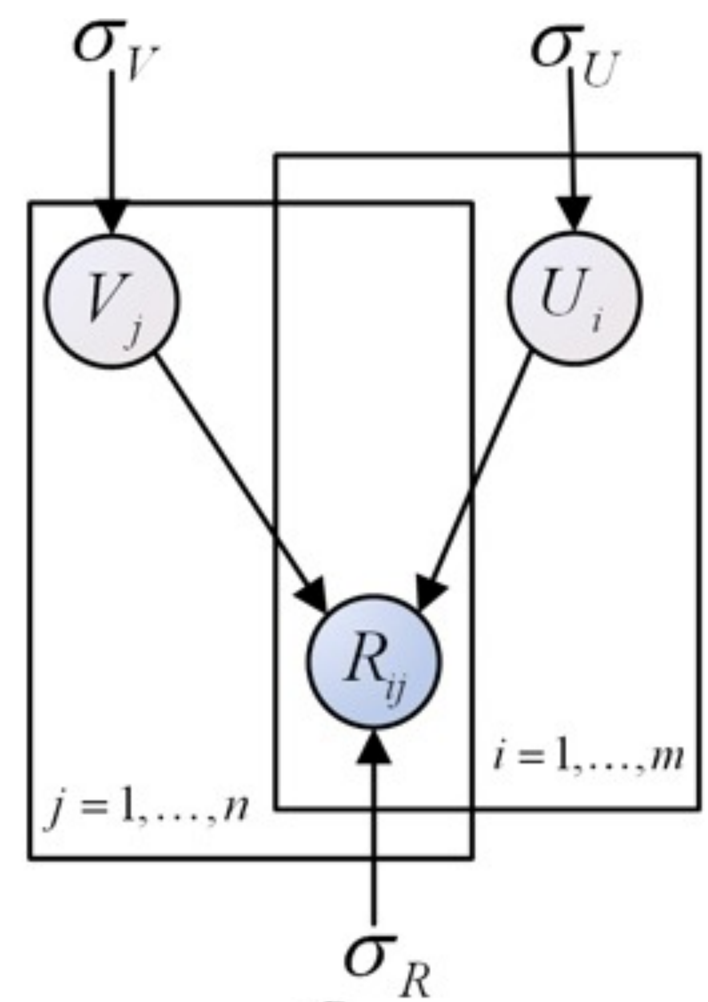
2nd Motivation

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



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



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

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

[R. Salakhutdinov, et al., NIPS2008]

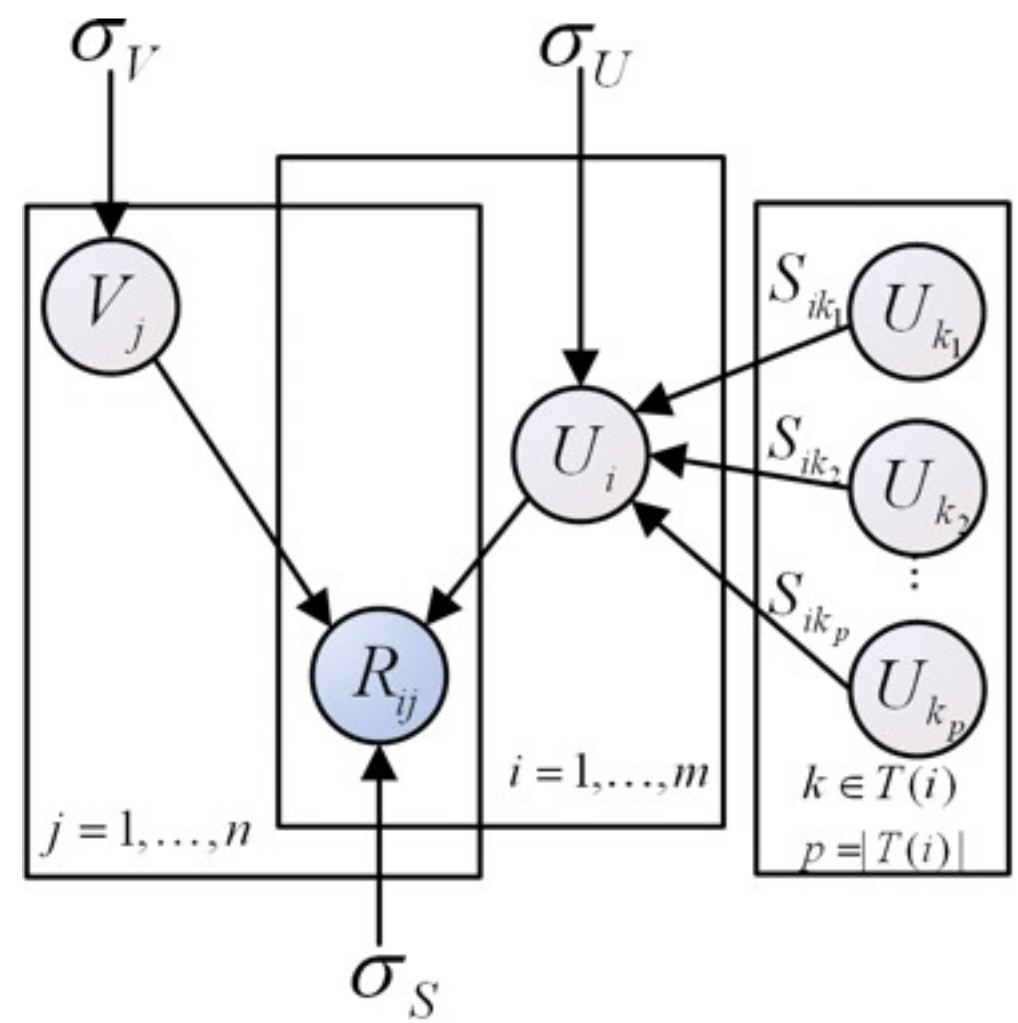
Recommendations by Trusted Friends

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

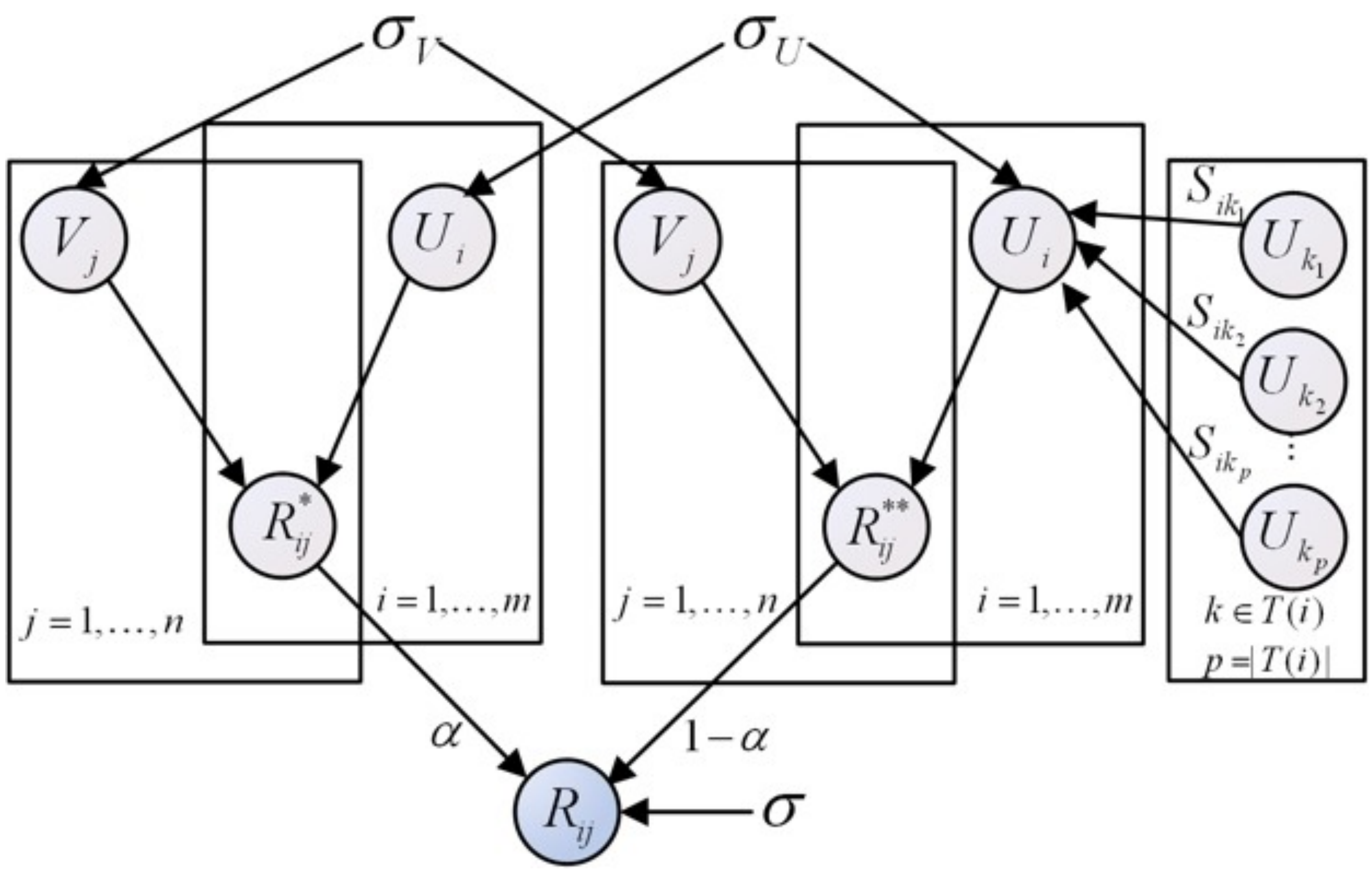
$$\widehat{R}_{ik} = \sum_{j\in\mathcal{T}(i)} R_{jk}S_{ij}$$

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

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



Recommendation with Social Trust Ensemble



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

Recommendation with Social Trust Ensemble

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

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

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

Complexity

- In general, the complexity of this method is linear with the observations the user-item matrix

Epinions Dataset

- 51,670 users who rated 83,509 items with totally 631,064 ratings
- Rating Density 0.015%
- The total number of issued trust statements is 511,799

Metrics

- Mean Absolute Error and Root Mean Square Error

$$MAE = \frac{\sum_{i,j} |r_{i,j} - \widehat{r}_{i,j}|}{N}$$

$$RMSE = \sqrt{\frac{\sum_{i,j} (r_{i,j} - \widehat{r}_{i,j})^2}{N}}$$

Comparisons

Table 3: Performance Comparisons (A Smaller MAE or RMSE Value Means a Better Performance)

Training Data	Metrics	Dimensionality = 5				Dimensionality = 10			
		Trust	PMF	SoRec	RSTE	Trust	PMF	SoRec	RSTE
90%	MAE	0.9054	0.8676	0.8484	**0.8377**	0.9039	0.8651	0.8426	**0.8367**
	RMSE	1.1959	1.1575	1.1418	**1.1109**	1.1917	1.1544	1.1365	**1.1094**
80%	MAE	0.9221	0.8951	0.8654	**0.8594**	0.9215	0.8886	0.8605	**0.8537**
	RMSE	1.2140	1.1826	1.1517	**1.1346**	1.2132	1.1760	1.1586	**1.1256**

PMF --- R. Salakhutdinov and A. Mnih (NIPS 2008)

SoRec --- H. Ma, H. Yang, M. R. Lyu and I. King (CIKM 2008)

Trust, RSTE --- H. Ma, I. King and M. R. Lyu (SIGIR 2009)

Comparisons

Table III: Performance Comparisons (A Smaller MAE or RMSE Value Means a Better Performance)

Training Data	Metrics	Dimensionality = 5						
		UserMean	ItemMean	NMF	PMF	Trust	SoRec	RSTE
90%	MAE	0.9134	0.9768	0.8738	0.8676	0.9054	0.8442	**0.8377**
	RMSE	1.1688	1.2375	1.1649	1.1575	1.1959	1.1333	**1.1109**
80%	MAE	0.9285	0.9913	0.8975	0.8951	0.9221	0.8638	**0.8594**
	RMSE	1.1817	1.2584	1.1861	1.1826	1.2140	1.1530	**1.1346**

Training Data	Metrics	Dimensionality = 10						
		UserMean	ItemMean	NMF	PMF	Trust	SoRec	RSTE
90%	MAE	0.9134	0.9768	0.8712	0.8651	0.9039	0.8404	**0.8367**
	RMSE	1.1688	1.2375	1.1621	1.1544	1.1917	1.1293	**1.1094**
80%	MAE	0.9285	0.9913	0.8951	0.8886	0.9215	0.8580	**0.8537**
	RMSE	1.1817	1.2584	1.1832	1.1760	1.2132	1.1492	**1.1256**

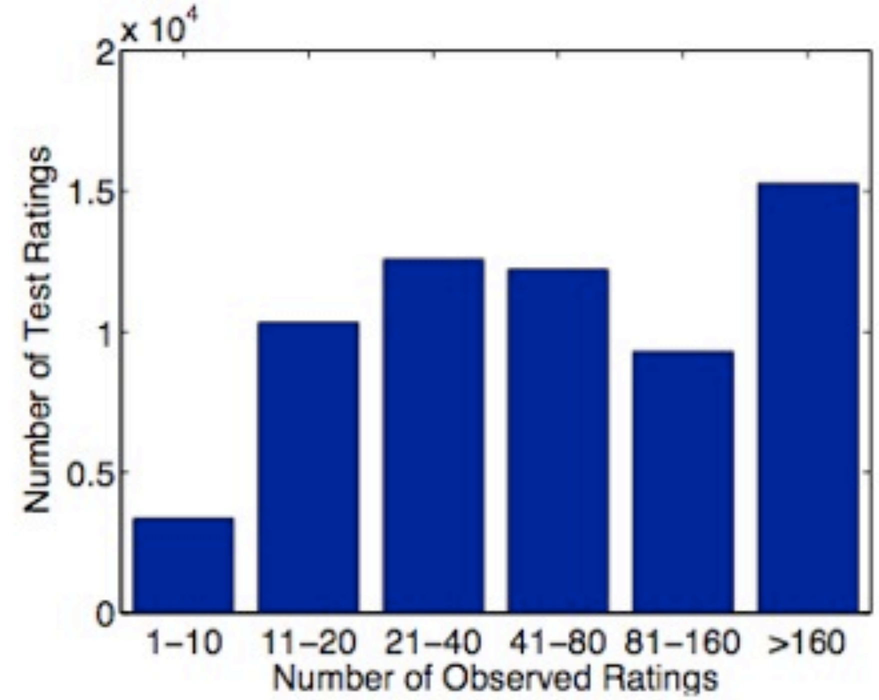
NMF --- D. D. Lee and H. S. Seung (Nature 1999)
PMF --- R. Salakhutdinov and A. Mnih (NIPS 2008)
SoRec --- H. Ma, H. Yang, M. R. Lyu and I. King (CIKM 2008)
Trust, RSTE --- H. Ma, I. King and M. R. Lyu (SIGIR 2009)

Performance on Different Users

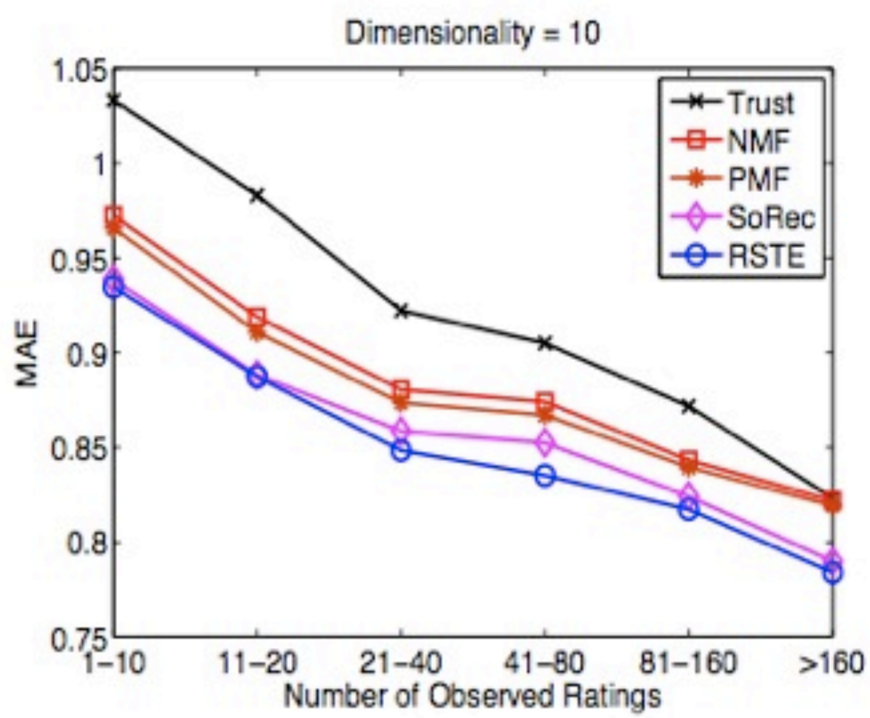
- Group all the users based on the number of observed ratings in the training data

- 6 classes: "1 − 10", "11 − 20", "21 − 40", "41 − 80", "81 − 160", "> 160",

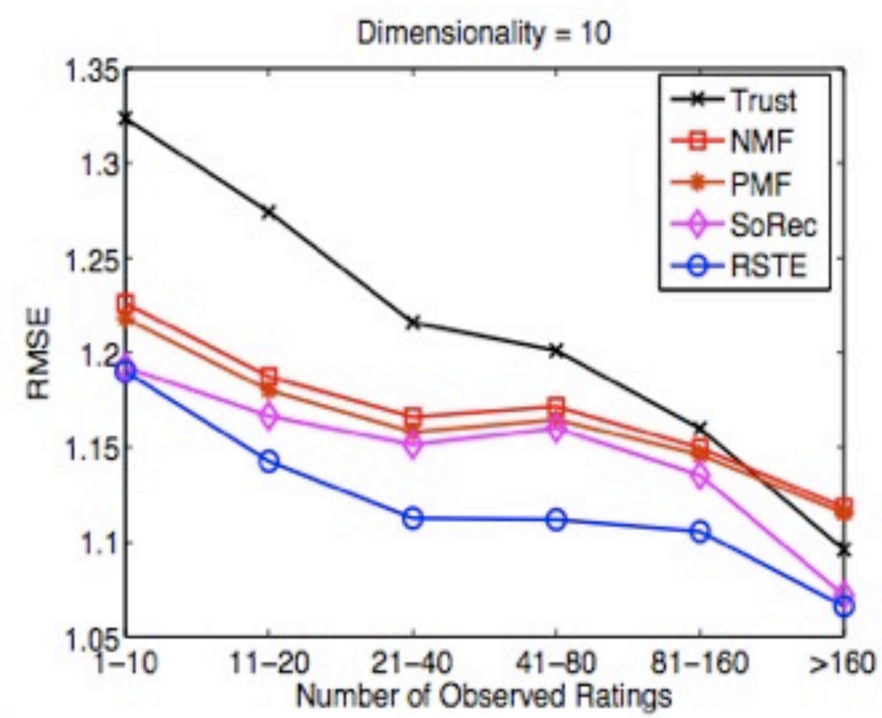
Performance on Different Users



(a) Distribution of Testing Data (90% as Training Data)

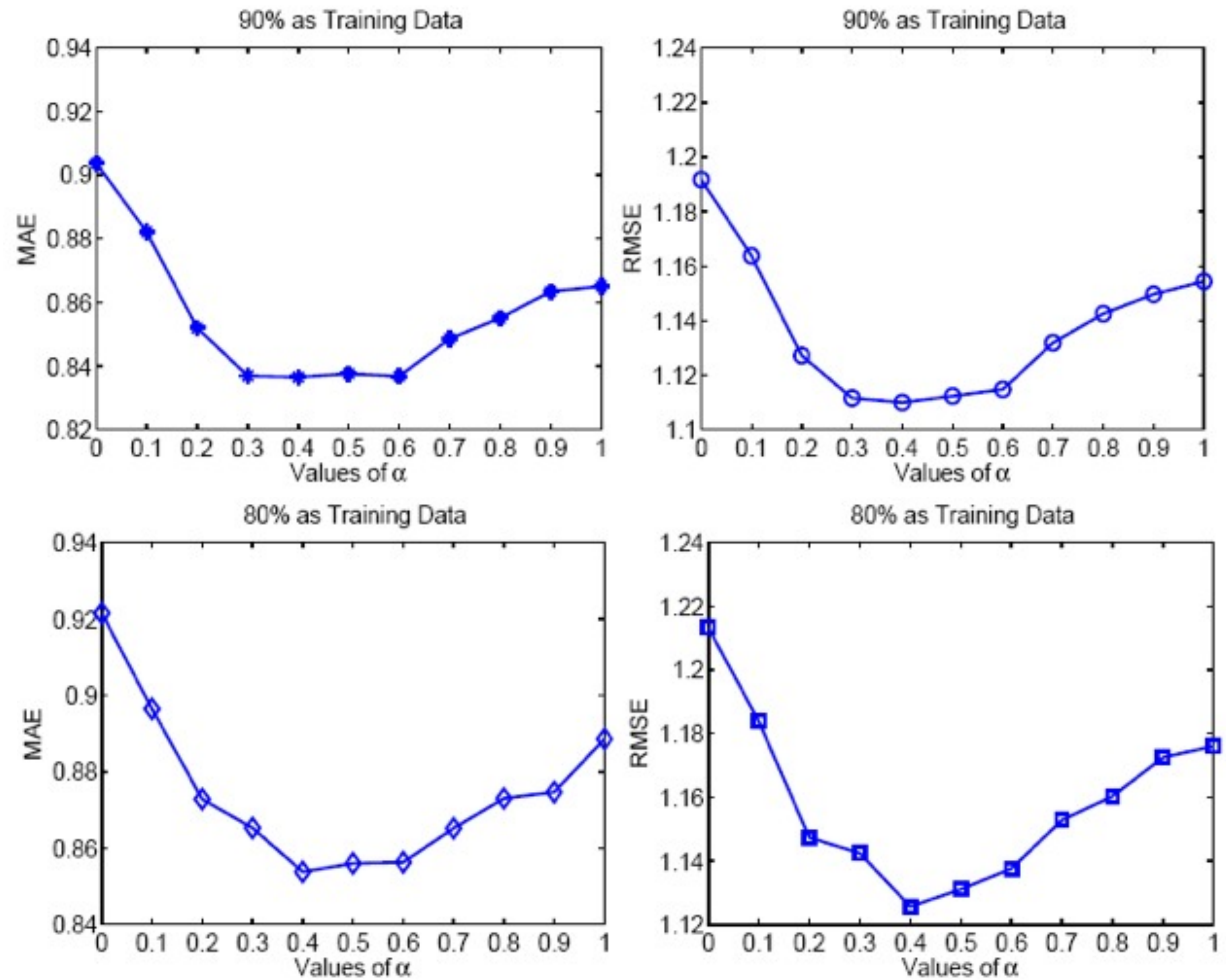


(b) MAE Comparison on Different User Rating Scales (90% as Training Data)



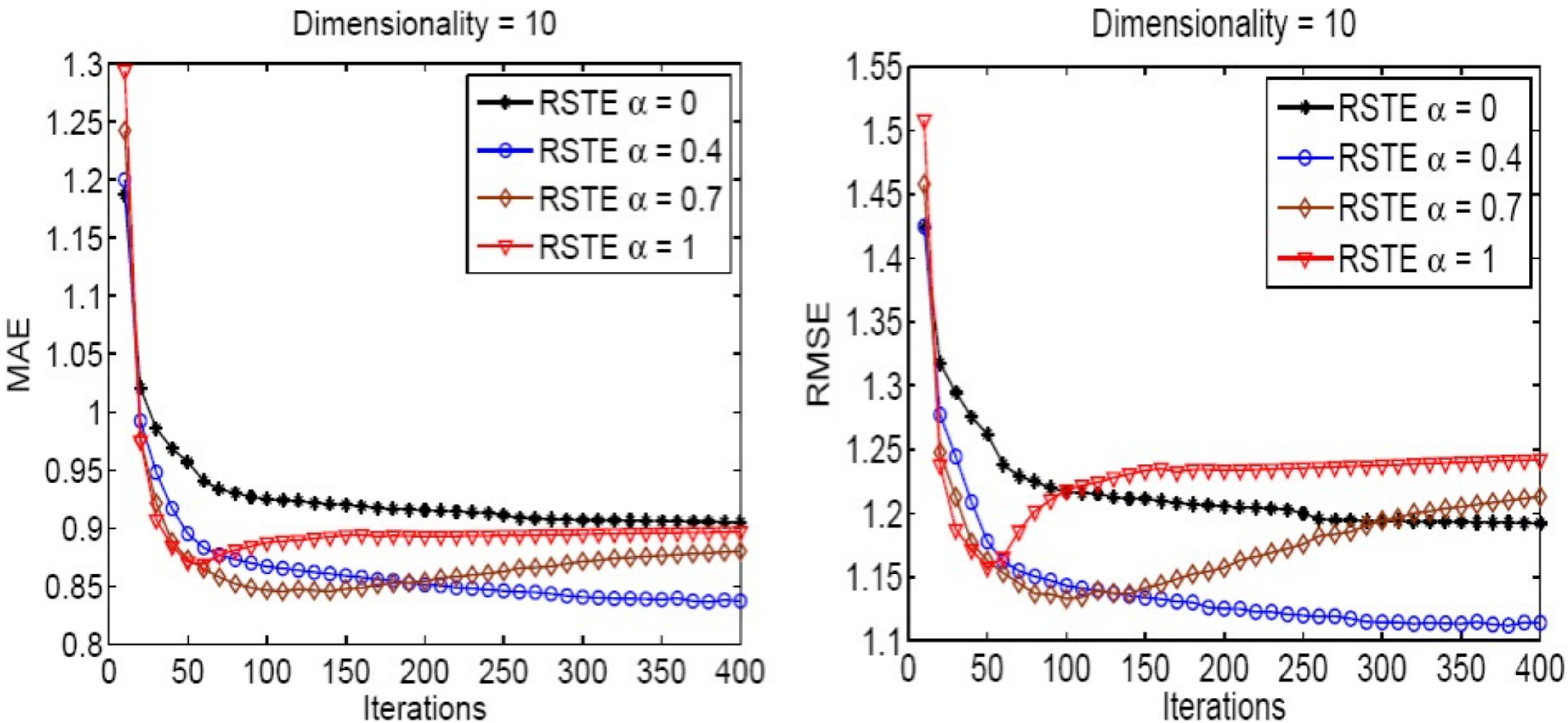
(c) RMSE Comparison on Different User Rating Scales (90% as Training Data)

Impact of Parameter Alpha



Impact of Parameter α (Dimensionality = 10)

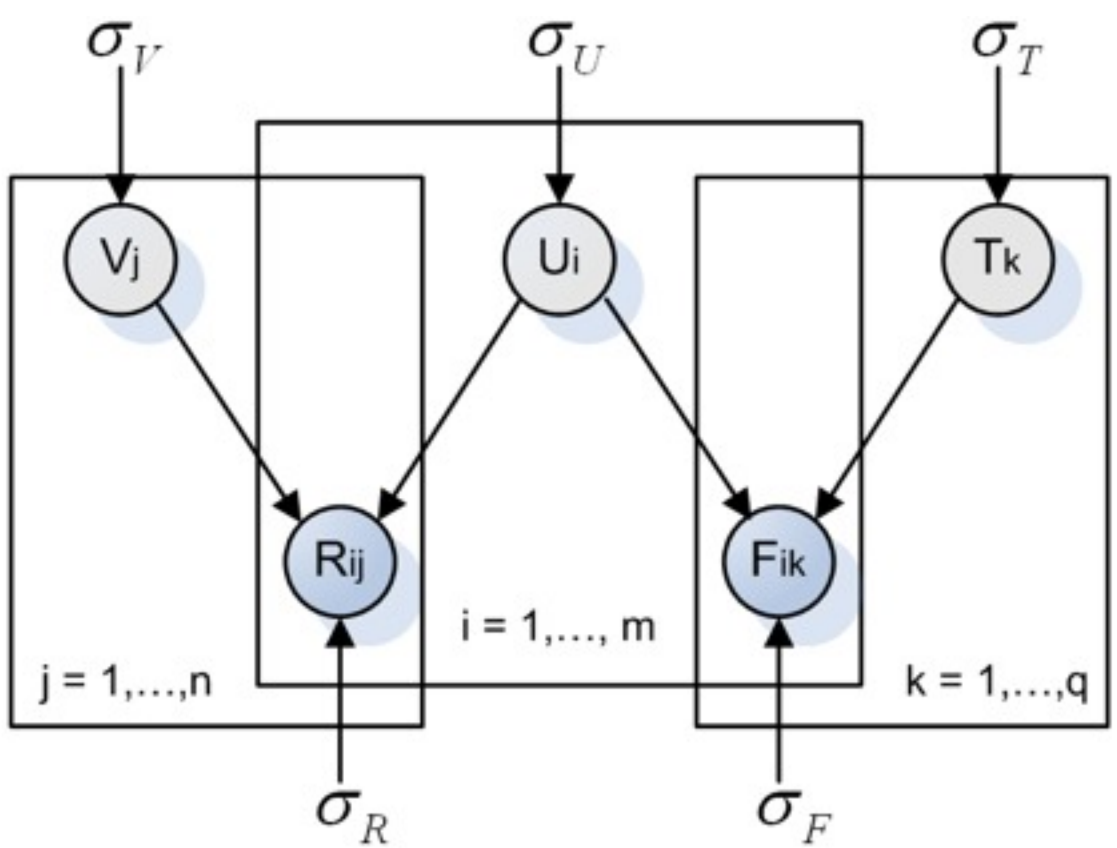
MAE and RMSE Changes with Iterations

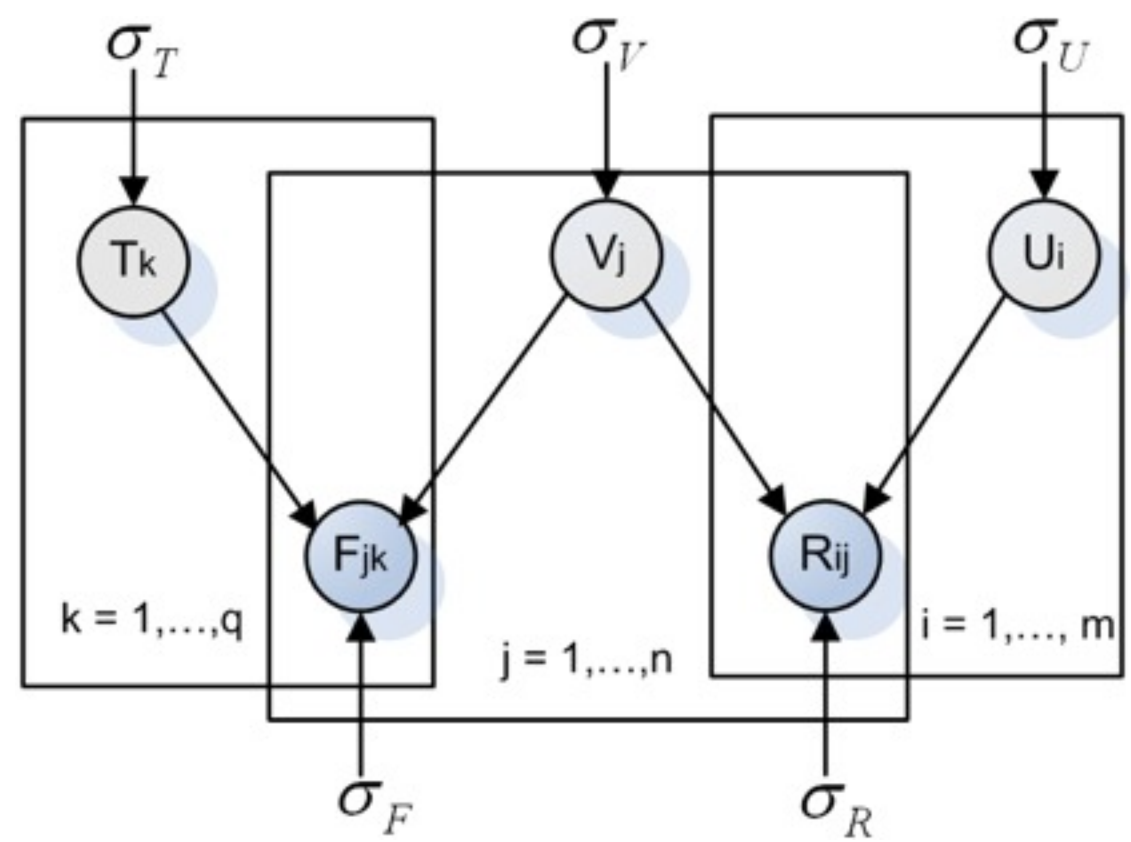


90% as Training Data

Further Discussion of SoRec

- Improving Recommender Systems Using Social Tags





MovieLens Dataset

71,567 users, 10,681 movies,

10,000,054 ratings, 95,580 tags

Further Discussion of SoRec

- MAE

Table V: MAE comparison with other approaches on MovieLens dataset
(A smaller MAE value means a better performance)

Methods		80% Training	50% Training	30% Training	10% Training
User Mean		0.7686	0.7710	0.7742	0.8234
Item Mean		0.7379	0.7389	0.7399	0.7484
5D	SVD	0.6390	0.6547	0.6707	0.7448
	PMF	0.6325	0.6542	0.6698	0.7430
	SoRecUser	0.6209	0.6419	0.6607	0.7040
	SoRecItem	**0.6199**	**0.6407**	**0.6395**	**0.7026**
10D	SVD	0.6386	0.6534	0.6693	0.7431
	PMF	0.6312	0.6530	0.6683	0.7417
	SoRecUser	0.6197	0.6408	0.6595	0.7028
	SoRecItem	**0.6187**	**0.6395**	**0.6584**	**0.7016**

Further Discussion of SoRec

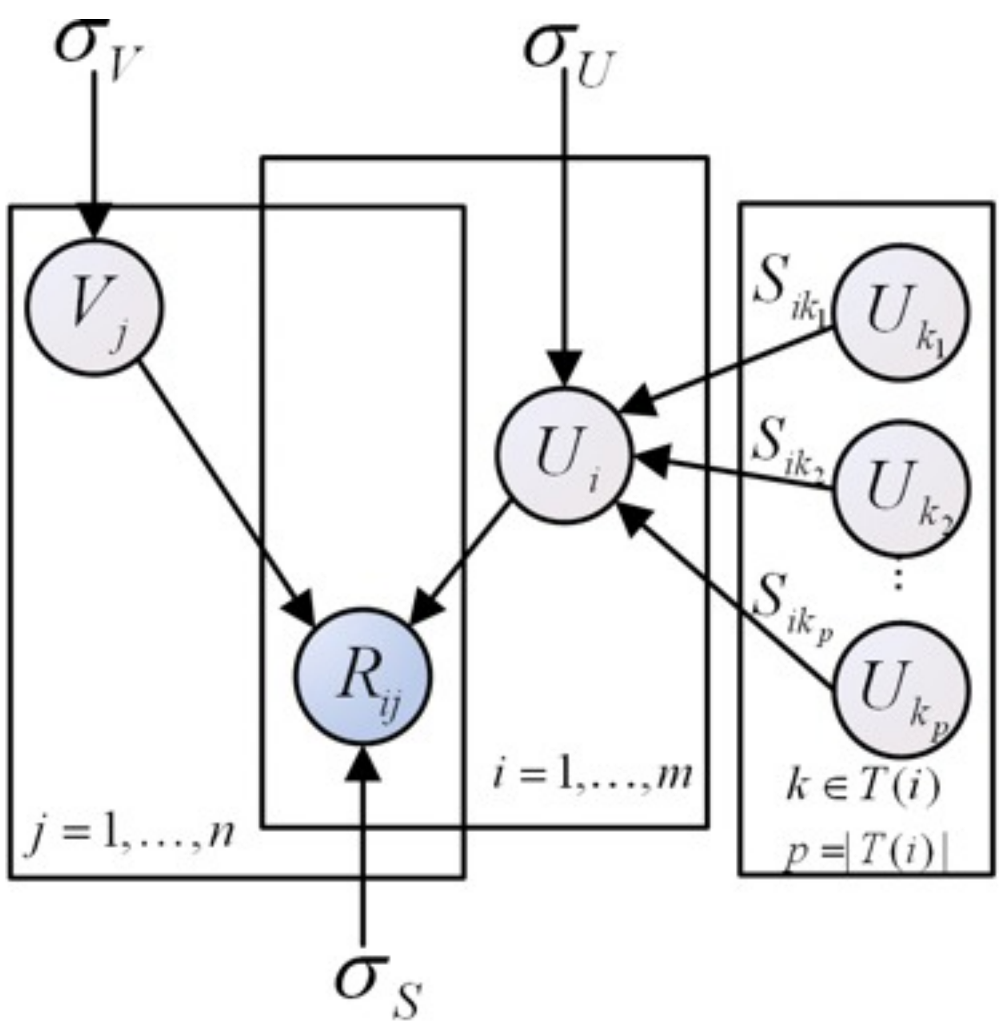
- RMSE

Table VI: RMSE comparison with other approaches on MovieLens dataset (A smaller RMSE value means a better performance)

Methods		80% Training	50% Training	30% Training	10% Training
User Mean		0.9779	0.9816	0.9869	1.1587
Item Mean		0.9440	0.9463	0.9505	0.9851
5D	SVD	0.8327	0.8524	0.8743	0.9892
	PMF	0.8310	0.8582	0.8758	0.9698
	SoRecUser	0.8121	0.8384	0.8604	0.9042
	SoRecItem	**0.8112**	**0.8370**	**0.8591**	**0.9033**
10D	SVD	0.8312	0.8509	0.8728	0.9878
	PMF	0.8295	0.8569	0.8743	0.9681
	SoRecUser	0.8110	0.8372	0.8593	0.9034
	SoRecItem	**0.8097**	**0.8359**	**0.8578**	**0.9019**

Further Discussion of RSTE

- Relationship with Neighborhood-based methods



- The trusted friends are actually the explicit neighbors
- We can easily apply this method to include implicit neighbors
- Using PCC to calculate similar users for every user

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
  - Many 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.

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

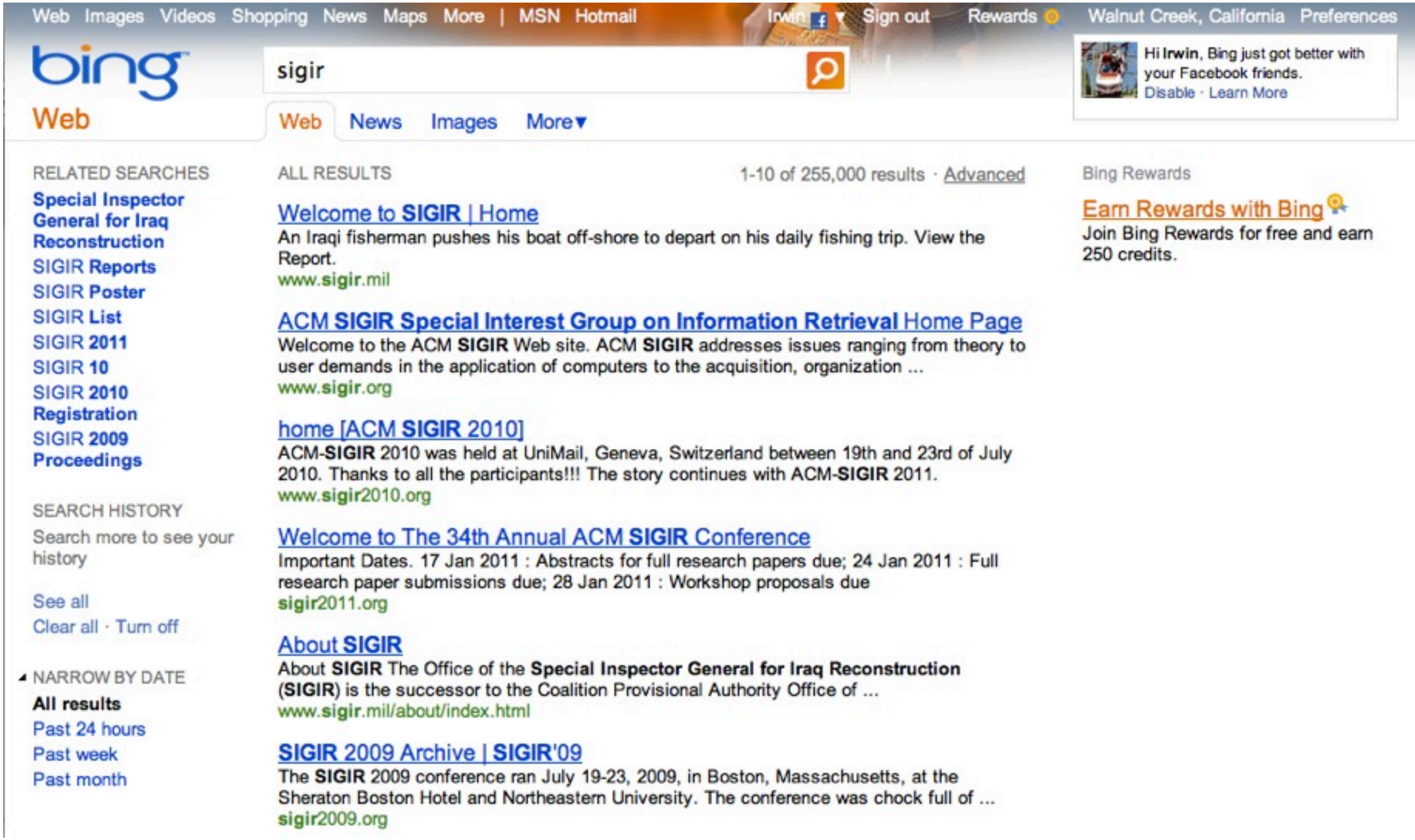
$$\min_{U} \frac{1}{2} \sum_{i=1}^{m} \sum_{t \in \mathcal{T}^+(i)} S_{it}^{\mathcal{T}} \|U_i - U_t\|_F^2$$

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

Web Site Recommendation

[Ma et al., SIGIR 2011]

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

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

Web users	v_1	v_2	v_3	v_4	v_5	v_6
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

Web users	z_1	z_2	z_3	z_4	z_5
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

Matrix Factorization

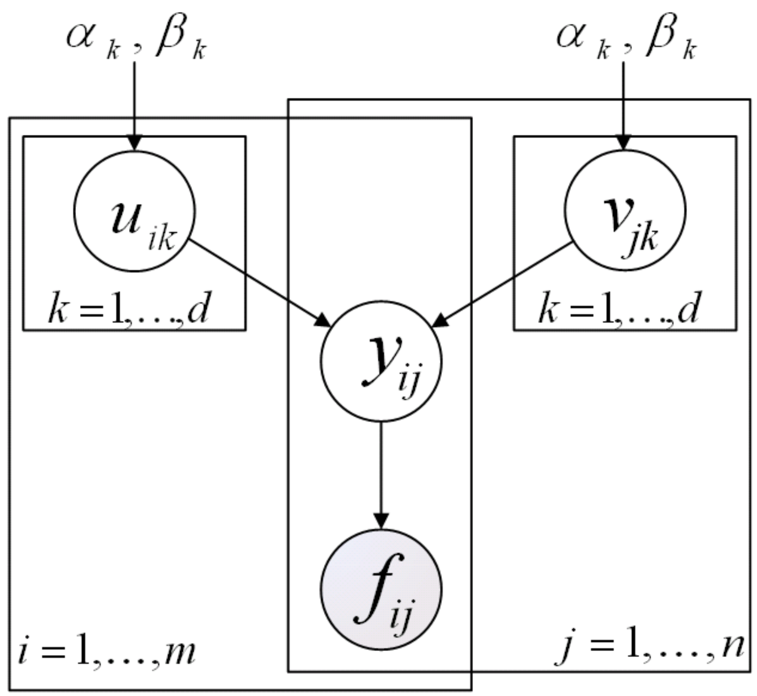
$$\min_{U,V} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}(R_{ij} - U_i V_j^T)^2 + \frac{\lambda_1}{2} ||U||_F^2 + \frac{\lambda_2}{2} ||V||_F^2$$

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

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

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

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

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

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

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

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

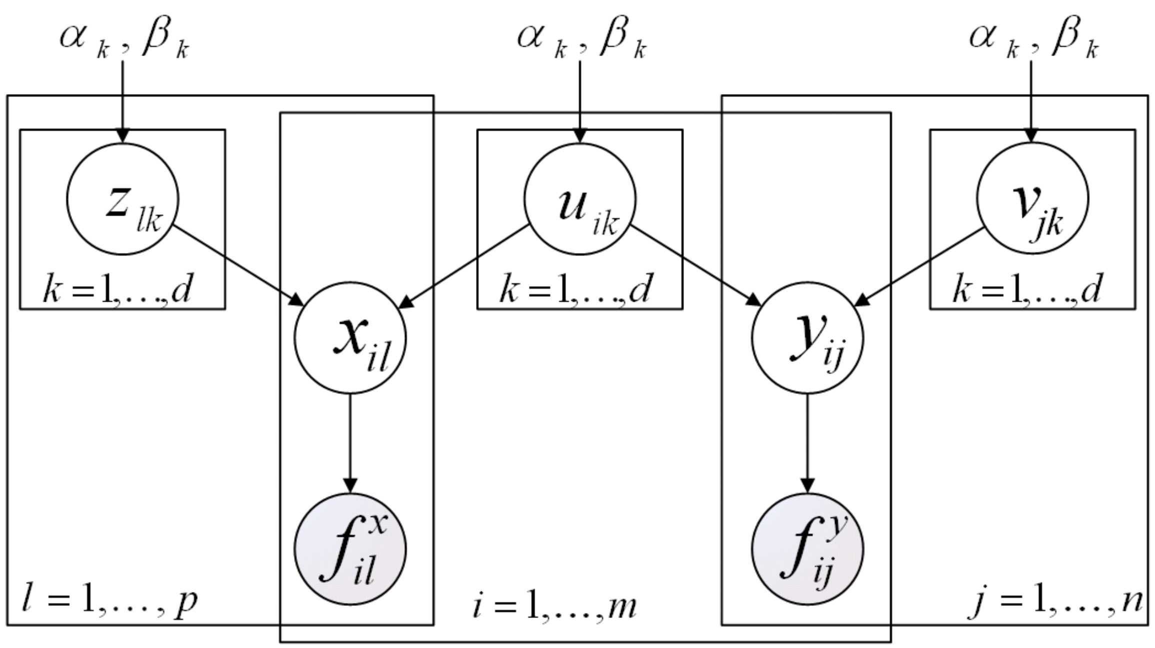
$$\begin{aligned} \mathcal{L}(U, V; F) \quad = \quad & \sum_{i=1}^{m} \sum_{k=1}^{d} ((\alpha_k - 1) \ln(u_{ik}/\beta_k) - u_{ik}/\beta_k) \\ & + \sum_{j=1}^{n} \sum_{k=1}^{d} ((\alpha_k - 1) \ln(v_{jk}/\beta_k) - v_{jk}/\beta_k) \\ & + \sum_{i=1}^{m} \sum_{j=1}^{n} (f_{ij} \ln y_{ij} - y_{ij}) + \text{const.}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

Probabilistic Factor Model

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

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

Collective Probabilistic Factor Model



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

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

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

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

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

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
- In user-site frequency matrix 2,612,016 entries, while in user-query frequency matrix 833,581 entries

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

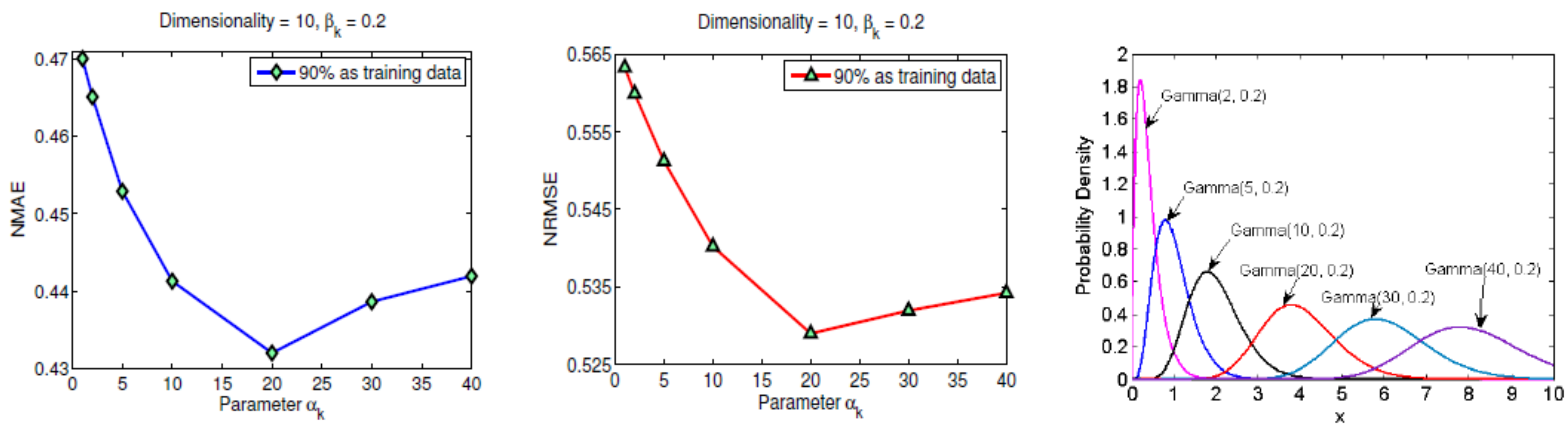
Statistics	User-Site Frequency	User-Query Frequency
Min. Num.	4	10
Max. Num.	9,969	4,693
Avg. Num.	20.33	23.05

Performance Comparison

Table 3: Performance Comparison (Dimensionality = 10)

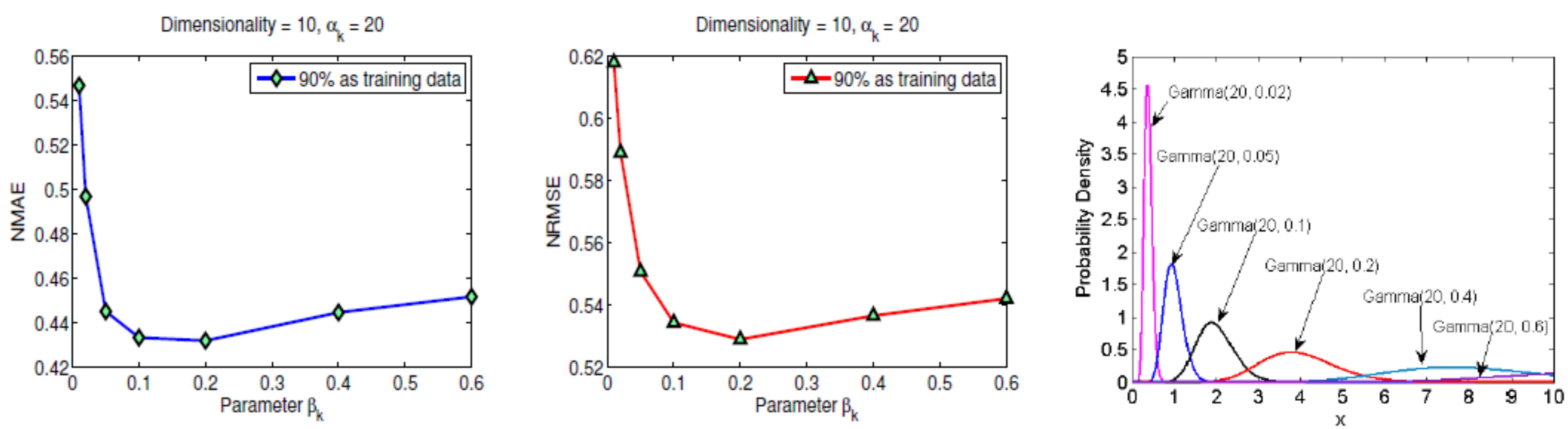
Training Data	Metrics	UserMean	SiteMean	SVD	PMF	NMF	GaP	PFM	CPFM
90%	NMAE	2.246	1.094	0.488	0.476	0.465	0.440	0.432	**0.427**
	Improve	80.98%	60.96%	12.50%	10.29%	8.17%	2.95%		
	NRMSE	3.522	2.171	0.581	0.570	0.554	0.532	0.529	**0.520**
	Improve	85.24%	76.05%	10.50%	8.77%	6.14%	2.26%		
80%	NMAE	2.252	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%		
	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%		

Impact of Parameters



(a) NMAE (b) NRMSE (c) Gamma Distributions

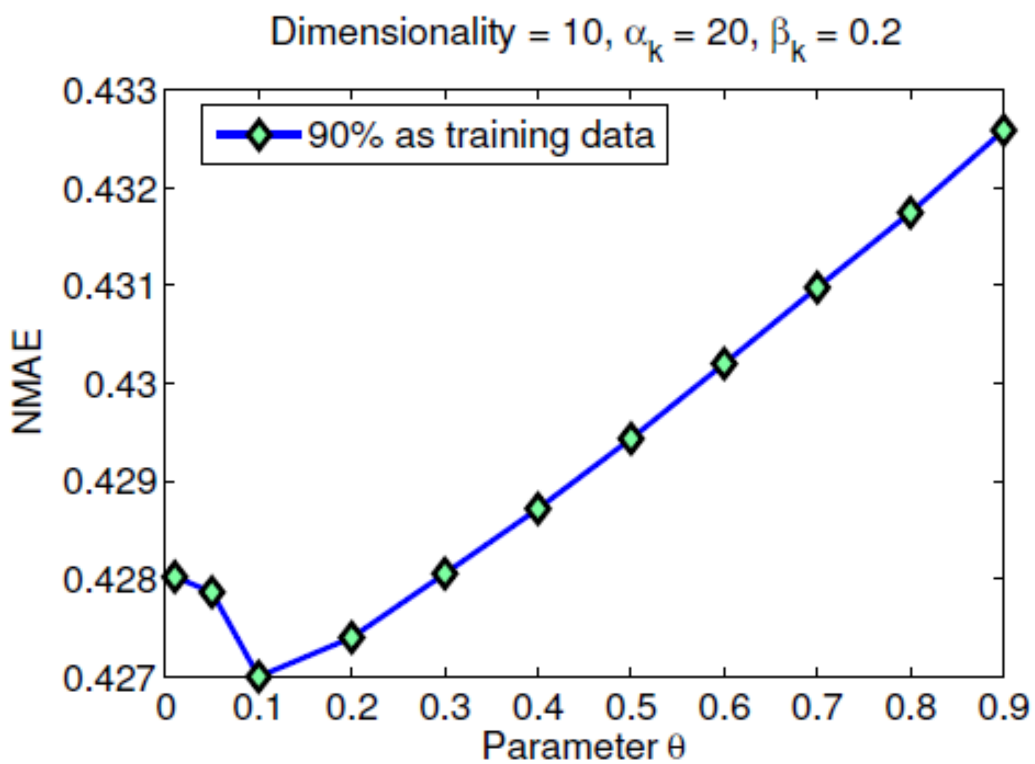
Figure 6: Impact of Parameter α_k in PFM



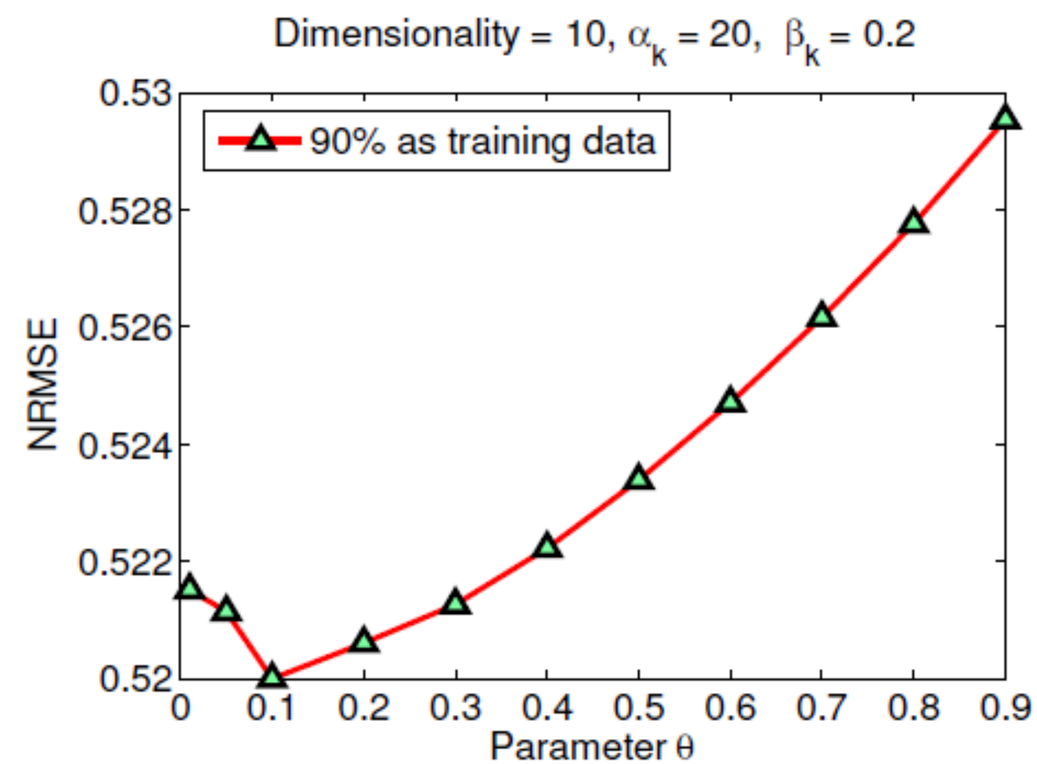
(a) NMAE (b) NRMSE (c) Gamma Distributions

Figure 7: Impact of Parameter β_k in PFM

Impact of Parameters



(a) NMAE



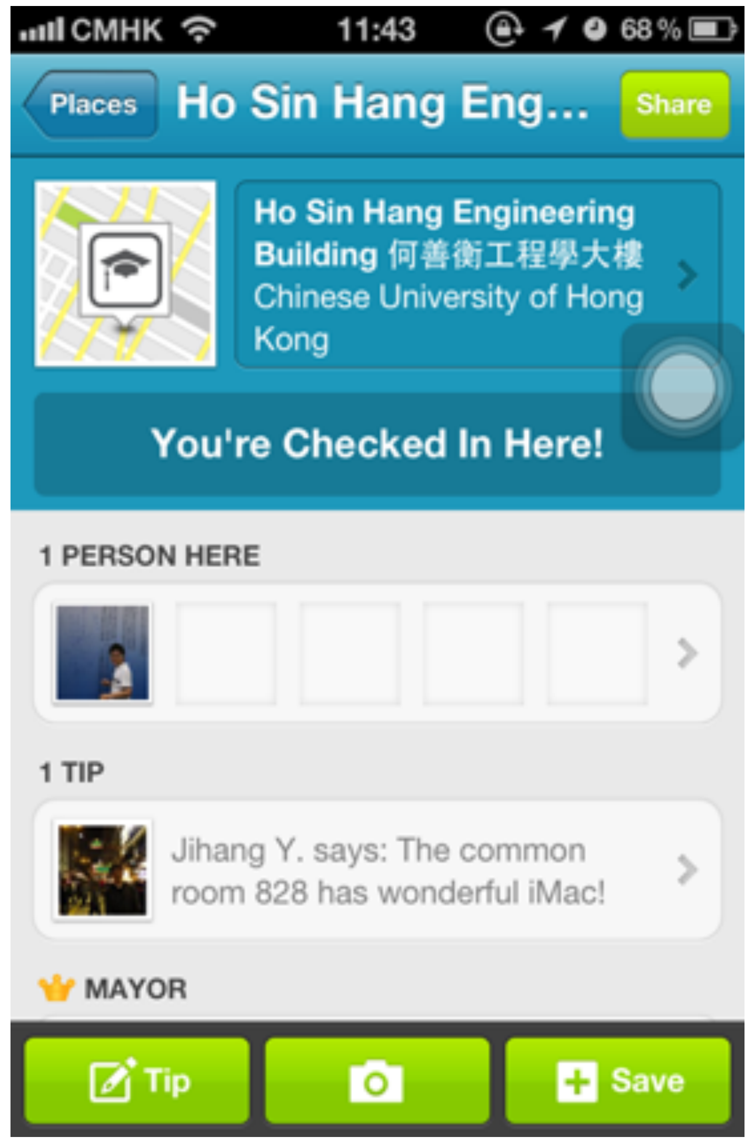
(b) NRMSE

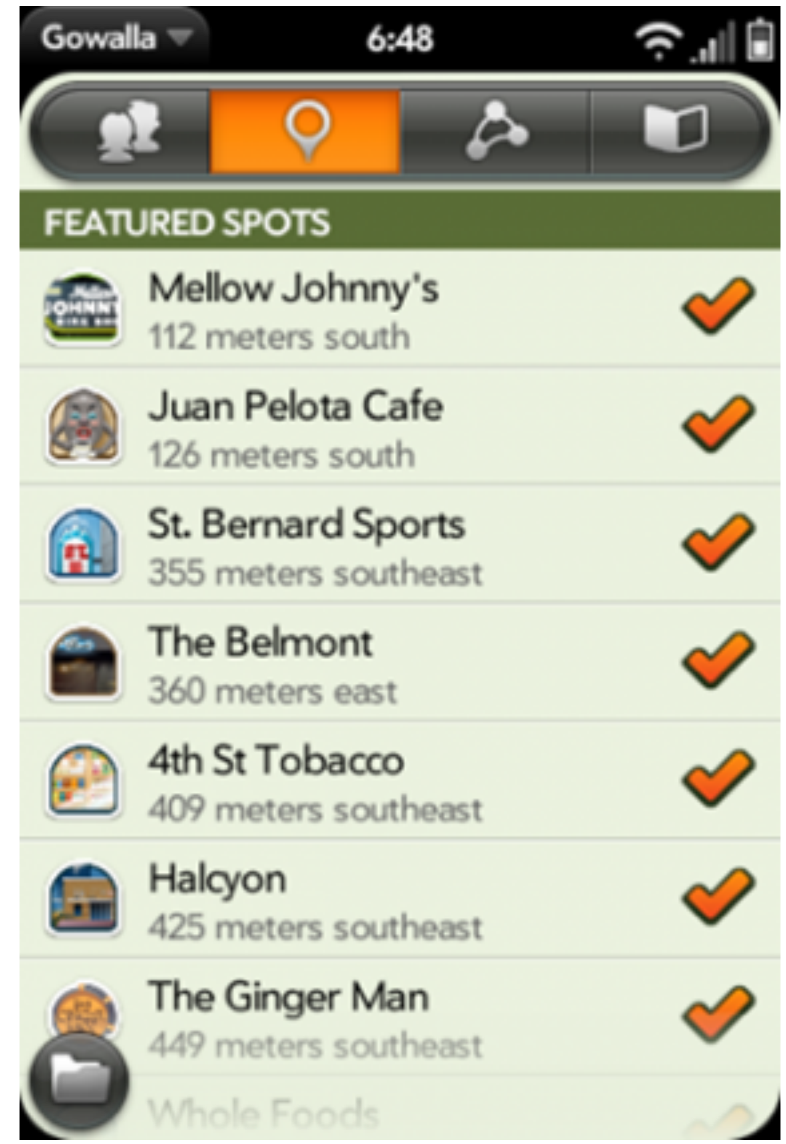
Figure 8: Impact of Parameter θ in CPFM

Location Recommendations

[Cheng et al., AAAI 2012]

Check Out on "Check-ins"





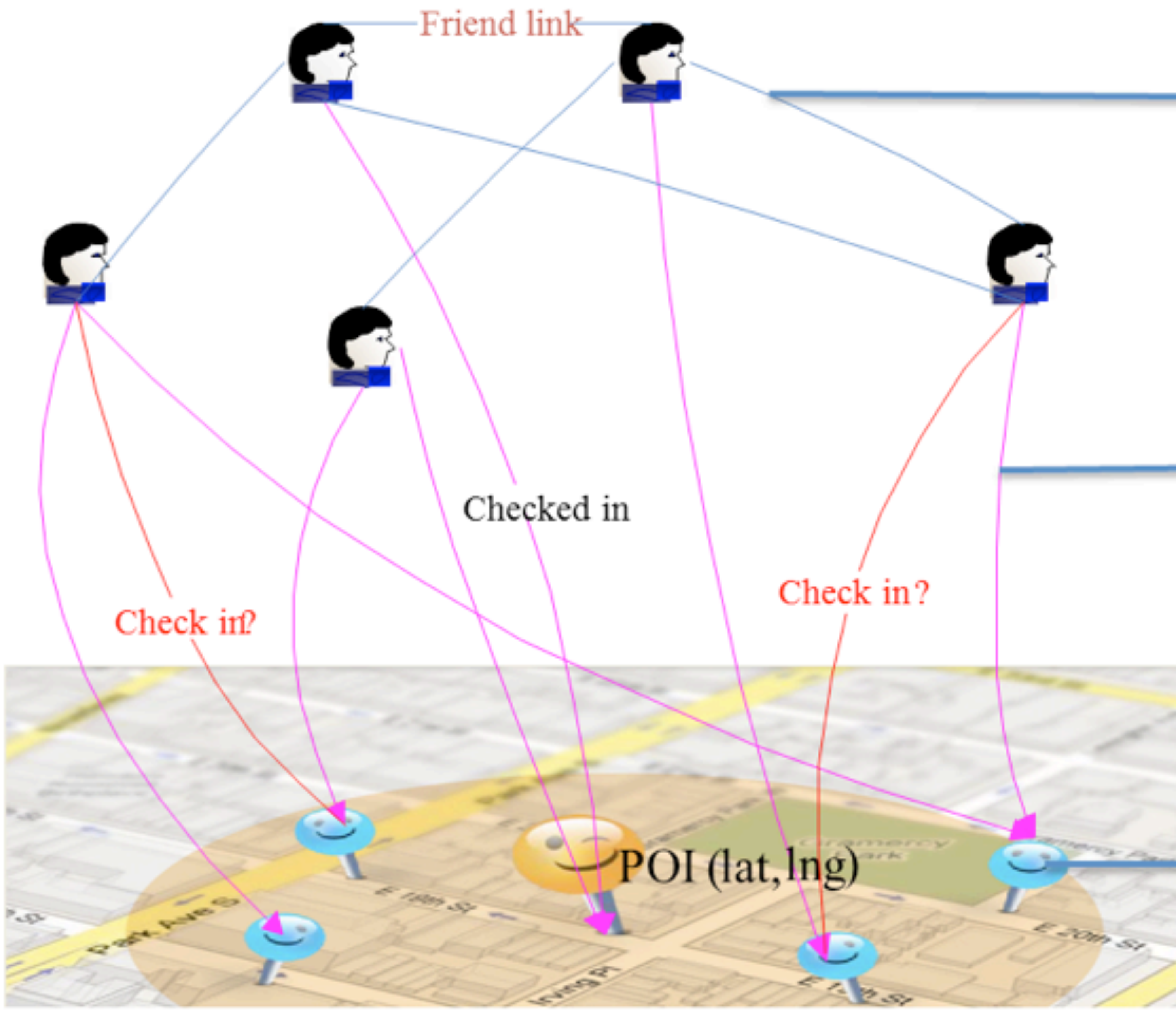
facebook

Facebook Places
Who. What. When. And now where.

foursquare

Gowalla

Location-based Social Networks (LBSNs)



- Community detection
- Link prediction

- POI recommendation
- Next place prediction

- Travel sequence detection
- Trip recommendation

Related Work

- POI recommendation on GPS trajectory logs
  - A collective matrix factorization method is applied on three matrices: location-activity, location-feature and activity-activity. [Zheng et al. 2010a]
  - A tensor factorization is conducted on the user-location-activity relationship. [Zheng et al. 2010b]
- POI recommendation on LBSNs dataset
  - A unified memory-based framework including user similarity, social and geographical influence, in which geographical influence in modeled as power-law distribution. [Ye et al. 2011]
  - Two-center mixture Gaussian model proposed to model human mobility in LBSNs. [Cho et al. 2011]

Motivations

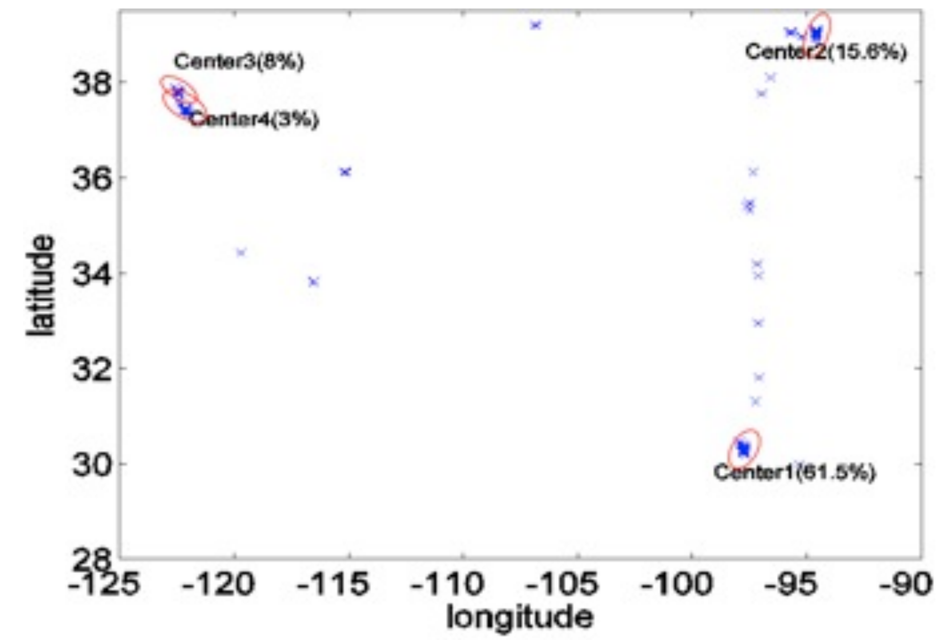
- Users have their personalized taste for different POIs.

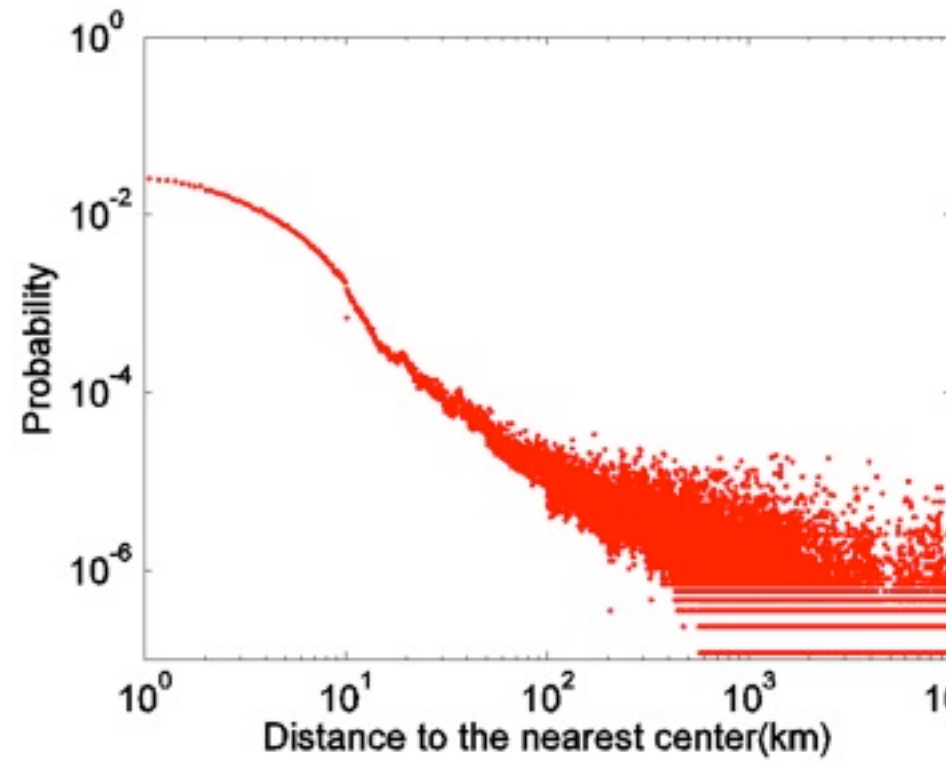



- The check-in probability is sensitive to geographical influence.

Observation #1

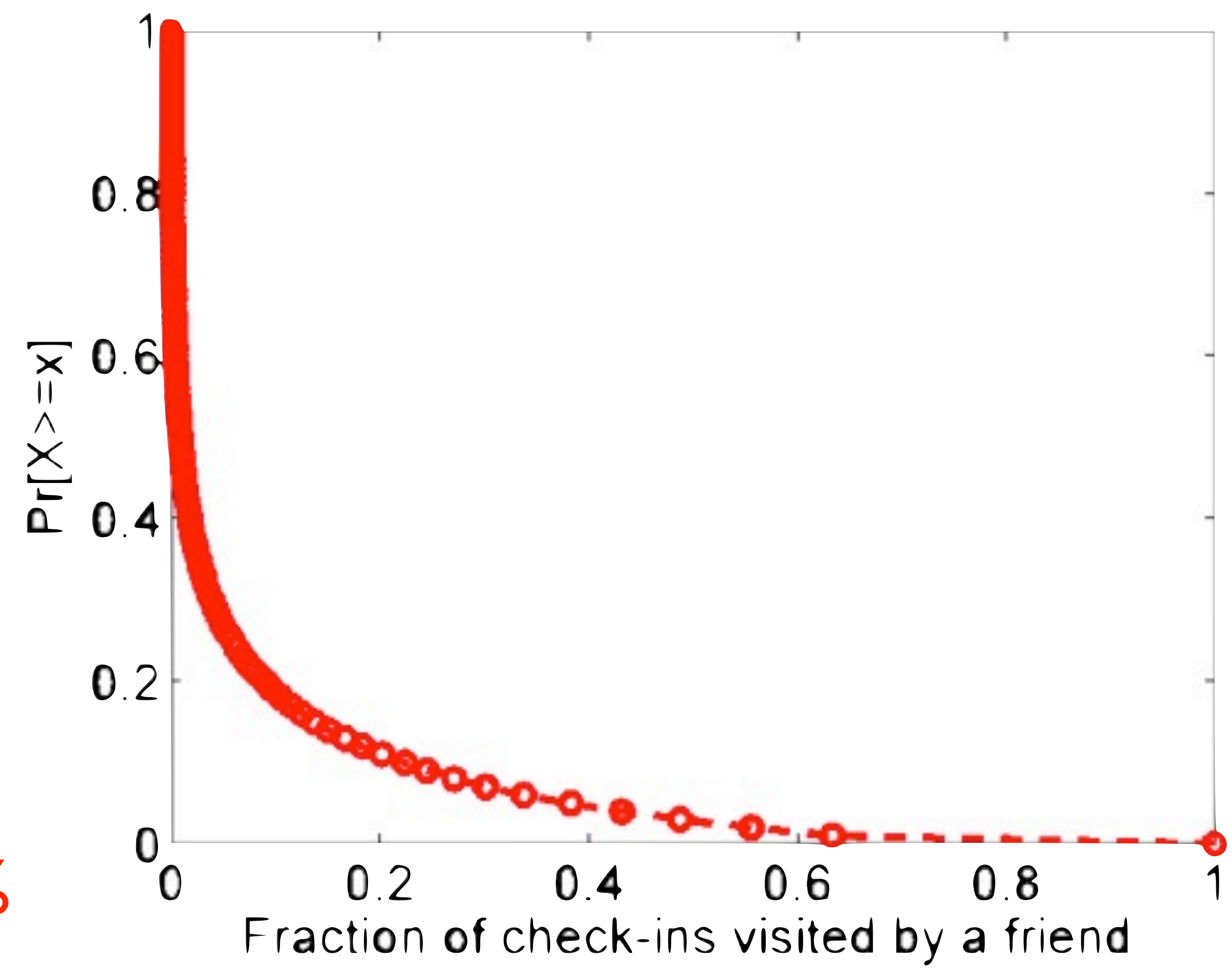
- Users tend to check-in around several centers
- Gaussian distribution to model check-ins at each center
- Inverse Distance Rule: check-in probability is inversely proportional to the distance to the nearest center





Observation #2

- Social information can help improve POI recommendation, but seems influence is limited
  - On average, overlap of a user's check-ins to his friends only about 9.6%
  - 90% users have only 20% common check-ins



Our Proposal

- Multi-center Gaussian Model (MGM) to capture geographical influence
- Propose a generalized fused matrix factorization framework to include social and geographical influences
- Experiments conducted on large-scale Gowalla dataset

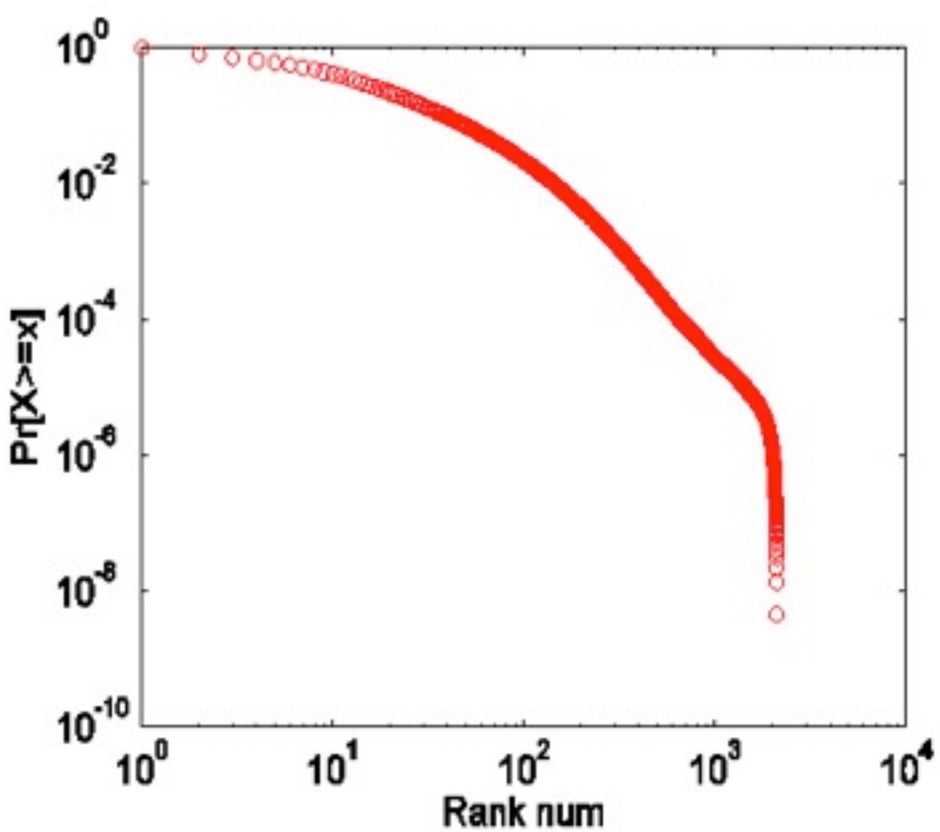
Multi-center Gaussian Model

- Notations
  - C_u: multi-center set for user u
  - f_{c_u}: total frequency at center c_u for user u
  - $\mathcal{N}(l|\mu_{c_u}, \Sigma_{c_u})$: the pdf of Gaussian distribution, μ_{c_u} and Σ_{c_u} denote the mean and covariance matrices of regions around center c_u
- The probability a user u visiting a location l given C_u is defined as:

$$P(l|C_u) = \sum_{c_u=1}^{|C_u|} \underbrace{P(l \in c_u)}_{\text{Which center?}} \underbrace{\frac{f_{c_u}^{\alpha}}{\sum_{i \in C_u} f_i^{\alpha}}}_{\text{Frequency normalization}} \underbrace{\frac{\mathcal{N}(l|\mu_{c_u}, \Sigma_{c_u})}{\sum_{i \in C_u} \mathcal{N}(l|\mu_i, \Sigma_i)}}_{\text{Gaussian distribution of the center}}.$$

Multi-center Discovery Algorithm

- A greedy clustering algorithm is proposed due to Pareto principle (top 20 locations cover about 80% check-ins)



for all user i in the user set $\mathcal{U}$ **do**
    Rank all check-in locations in $|\mathcal{L}|$ according to visiting frequency
    $\forall l_k \in L$, set $l_k.center = -1$;
    Center_list $= \emptyset$; center_no $= 0$;
    for $i = 1 \to |L|$ **do**
        if $l_i.center == -1$ **then**
            center_no++; Center $= \emptyset$; Center.total_freq $= 0$;
            Center.add(l_i); Center.total_freq += l_i.freq;
            for $j = i + 1 \to |L|$ **do**
                if $l_j.center == -1$ *and* $dist(l_i, l_j) \leq d$ **then**
                    $l_j.center$ = center_no; Center.add(l_j);
                    Center.total_freq += l_j.freq;
                end if
            end for
            if Center.total_freq $\geq |u_i|$.total_freq * θ **then**
                Center_list.add(Center);
            end if
        end if
    end for
    RETURN Center_list for user i;

Fused Framework

- Probabilistic Matrix Factorization (PMF) models users' **preference** on locations:$F \approx U^T L$, and the frequency will be converted to $[0, 1]$ by $g(x) = 1/(1 + \exp(-x))$.

- PFM with **Social Regularization** (PMFSR) [Ma et al. 2011b]:

$$
\begin{aligned}
\min_{U,L} \Omega(U, L) &= \sum_{i=1}^{|\mathcal{U}|} \sum_{j=1}^{|\mathcal{L}|} I_{ij}(F_{ij} - U_i^T L_j)^2 \\
&+ \beta \sum_{i=1}^{|\mathcal{U}|} \sum_{f \in \mathcal{F}(i)} Sim(i, f) ||U_i - U_f||_F^2 \\
&+ \lambda_1 ||U||_F^2 + \lambda_2 ||L||_F^2,
\end{aligned}
$$

- MGM models **geographical influence**

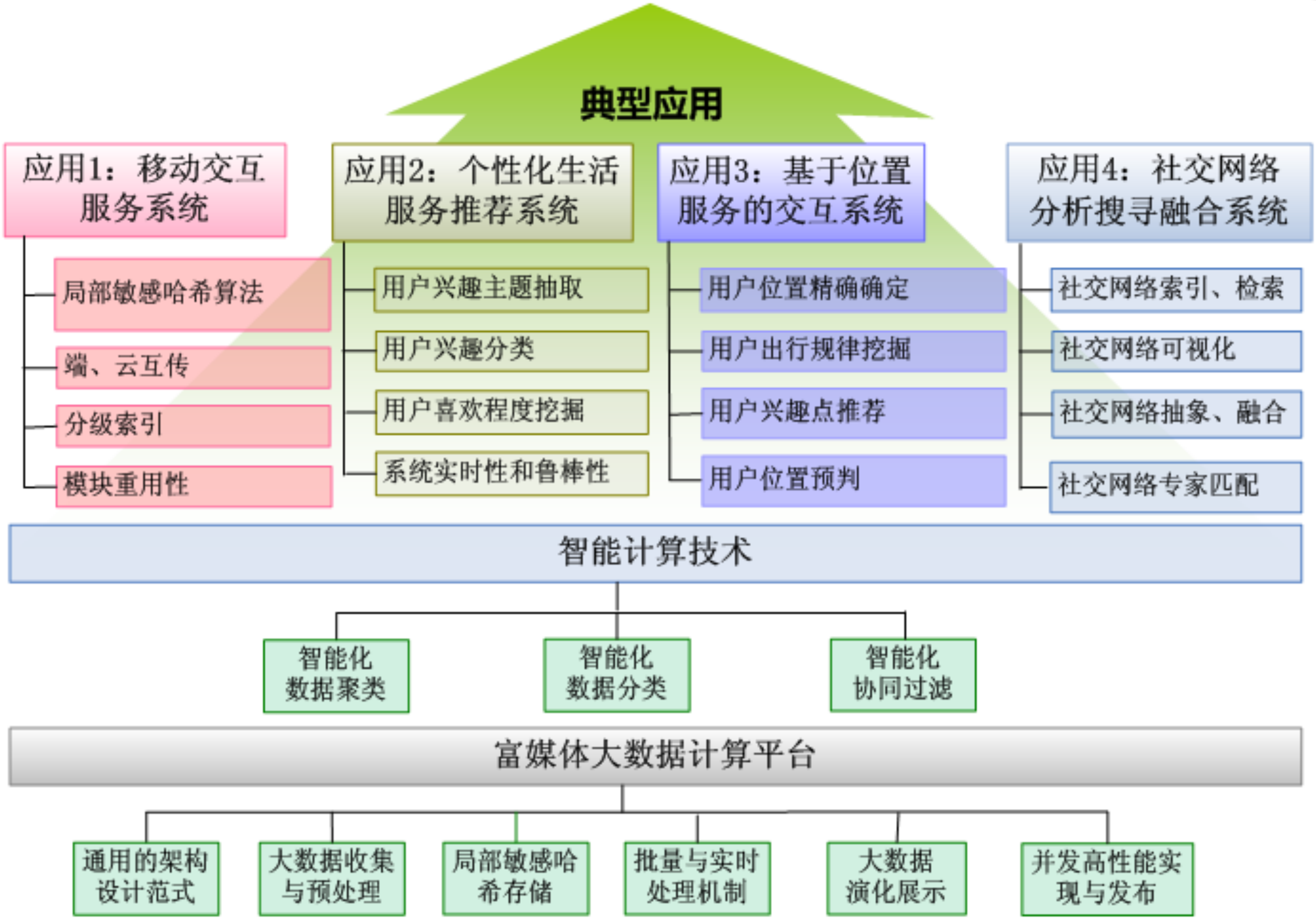
- We can fuse them together:

$$P_{ul} = \underbrace{\lambda P(F_{ul})}_{\text{Social Influence}} + \underbrace{(1 - \lambda) P(l|C_u)}_{\text{Geographical Influence}}, \text{ where } P(F_{ul}) \propto U_u^T L_l.$$

Concluding Remarks

- Both social and location recommendation play a significant role in the social web!
- Social recommendation extends traditional models and techniques by using social graphs, ensembles, distrust relationships, clicks, etc.
- Fusing of social behavior information, e.g., media consumption patters, temporal relationships, etc.
- Location recommendation follows a similar path with new data and features.

深圳市富媒体大数据重点实验室



Looking for People with Bright Ideas

- Postdoc
- PhD Students
- Engineers

- Machine learning, social computing, data mining/analytics, Big Data, etc.
  - Semi-supervised learning, on-line learning, active learning, multiple kernel learning, feature selection, matrix factorization, large margin classifiers, kernel methods, etc.
  - Recommender systems, Q&A, human computation, crowdsourcing, user profiling, user modeling, graph algorithms, computational advertising, etc.
  - Scalable algorithms, mobile apps, cloud computing, open sourced projects, etc.

The ACM Conference Series on
Recommender Systems

HOME | RECSYS 2013 | PAST CONFERENCES | BOOKS | WIKI | CONTACT

search...

ACM RecSys 2013

The ACM Recommender System conference is the premier international forum for the presentation of new research results, systems and techniques in the broad field of recommender systems. Recommendation is a particular form of information filtering, that exploits past behaviours and user similarities to generate a list of information items that is personally tailored to an end-user's preferences. The seventh conference in this series, RecSys 2013, will bring together researchers and practitioners from academia and industry to present the latest results and identify new trends and challenges in providing recommendation components in a range of innovative application contexts. As RecSys brings together the main international research groups working on recommender systems, along with many of the world's leading e-commerce companies, in the last number of years, it has become the most important annual conference for the presentation and discussion of recommender system research. In addition to the main technical track, RecSys 2013 program will feature keynote and invited talks, tutorials covering state-of-the-art in this domain, a workshop program, an industrial track and a doctoral symposium.

ACM RECSYS 2013 (HONG KONG)

- About the Conference
- Registration
- Important Dates
- Call for Contributions
- Program
- Keynotes
- Tutorials

CrowdRec 2013

Workshop on Crowdsourcing and Human Computation for Recommender Systems

Co-located with ACM RecSys 2013, Hong Kong, 12th October, 2013

Home	Submission	Committee	Programme

Announcements

2013-07-18
The deadline of submission has been extended to 26th July, 2013.

2013-05-24
CrowdRec 2013 Website launched.

Introduction

We are pleased to announce the first workshop on crowdsourcing and human computation for recommender systems. The recommender systems we refer to as a broad spectrum of applications involving recommendation, information valuation, filtering, summarization, etc. in various contexts from e-commerce to social networking and mobile applications. This workshop aims to provide a scholarly venue for researchers and practitioners to exchange the advances of crowdsourcing and human computation technologies and applications, with an emphasis on the applications in recommendation systems. The potentials and advantages of crowdsourcing and human computation have been explored for a number of areas such as computer-human interaction and information retrieval, we believe that these advances can also benefit the research of recommender systems at large.

Call For Papers

Important Dates

26th July, 2013 (PDT)
Submission due (Extended)

16th August, 2013
Author notification

30th August, 2013
Camera-ready version due

12th October, 2013
CrowdRec 2013 Workshop

Contacts

General Inquiries
Kuan-Ta Chen
swc@iis.sinica.edu.tw

Submission & Programme-related
Albert Au Yeung
albertauyeung@gmail.com

Overview

- Home
- Organization
- Contact
- Program

Partners

STC SOCIAL NETWORKING

+1 1

In conjunction with the 2013 IEEE International Conference on Big Data (IEEE Big Data 2013)

Workshop on Scalable Machine Learning: Theory and Applications

October 6, 2013, Santa Clara, CA, USA

Big Data are encountered in various areas, including Internet search, social networks, finance, business sectors, meteorology, genomics, connectomics, complex physics simulations, and biological and environmental research. The huge volume, high velocity, significant variety, and low veracity bring challenges to current machine learning techniques. It is desirable to scale up machine learning techniques for modeling and analyzing the big data from various domains.

The workshop aims to provide professionals, researchers, and technologists with a single forum where they can discuss and share the state-of-the-art of *scalable machine learning* technologies from theory and applications.

We thank the following experts for accepting our invitation to give plenary talks:

- Mikhail Bilenko, Microsoft research
- Carlos Guestrin, University of Washington
- Alek Kolcz, Twitter
- Alex Smola, Carnegie Mellon University

Topics of Interest

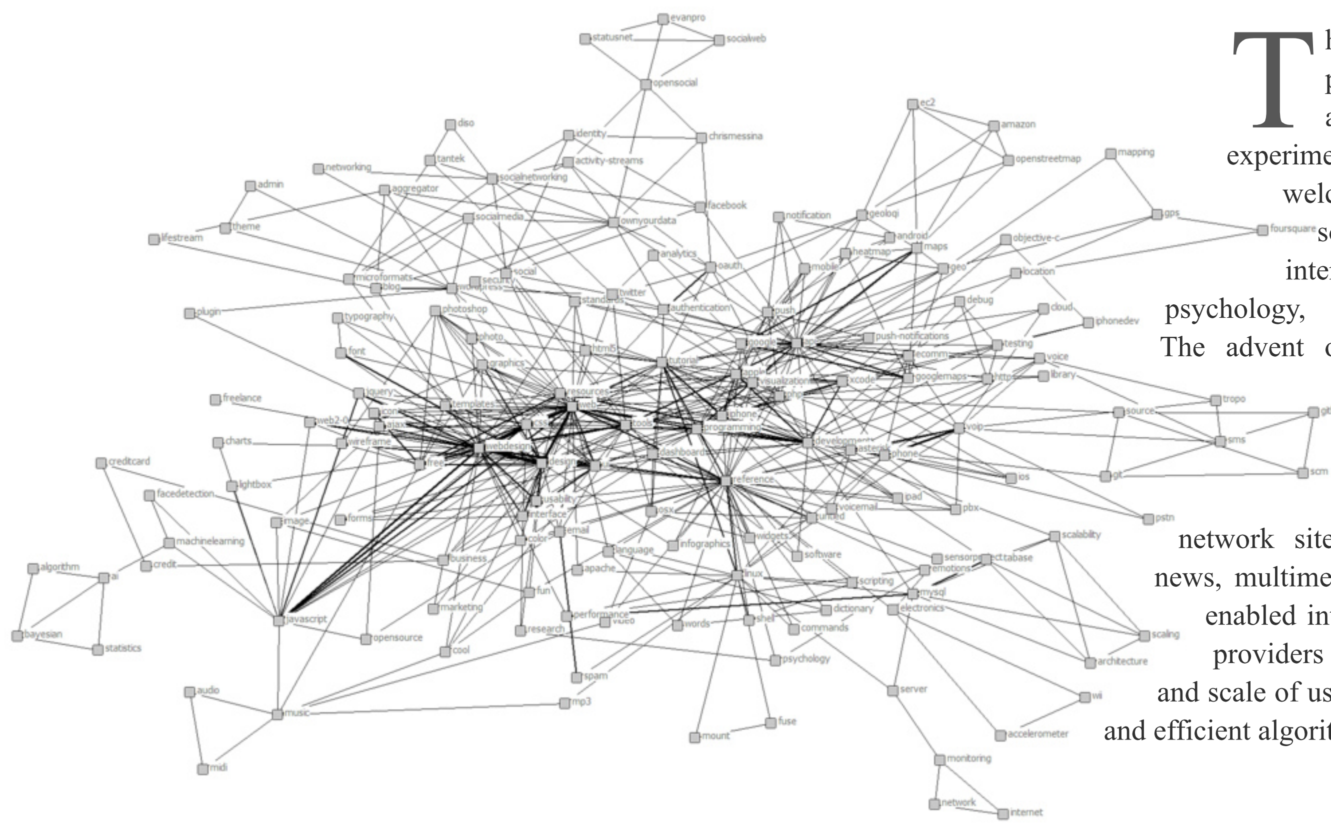
Topics of interest include, but not limited to:

- **Distributed machine learning architectures**
  - Data separation and integration techniques
  - Machine learning algorithms for GPUs
  - Machine learning algorithms for clouds
  - Machine learning algorithms for clusters

New Book Series

SOCIAL MEDIA & SOCIAL COMPUTING

CALL FOR BOOKS!



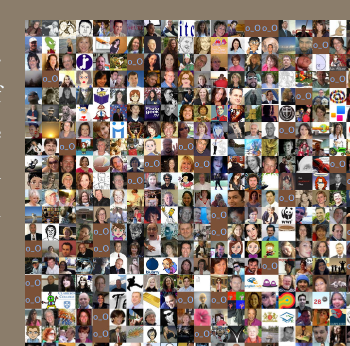
The Social Media and Social Computing Series focuses on publishing high quality references in the rapidly emerging area of social media and social computing. Both experimental/practical as well as theoretical investigations are welcome. The series targets both scholars and practitioners in social media and social computing for work in the intersection of computer science, information technology, psychology, economics, education and other social sciences. The advent of the Internet and the Web has resulted in social interactions and behaviors through the use of technologies and web services, e.g., hardware devices such as smart phones, tablets, RFID, etc., software services such as wikis, blogs, micro-blogs, social network sites, recommender systems, social bookmarking, social news, multimedia sharing sites, etc. Analyzing these technologically-enabled interactions in their social context will benefit information providers and information consumers. However, the large volume and scale of user-generated contents require effective modeling methods and efficient algorithms to handle these chalenging problems.

Series Editor:

Irwin King

Prof. King is Associate Editor of the IEEE Transactions on Neural Networks (TNN) and IEEE Computational Intelligence Magazine (CIM). He is a senior member of IEEE and a member of ACM, International Neural Network Society (INNS), and VP & Governing Board Member of the Asian Pacific Neural Network Assembly (APNNA) . He serves the Neural Network Technical Committee (NNTC) and the Data Mining Technical Committee under the IEEE Computational Intelligence Society.

https://www.cse.cuhk.edu.hk/irwin.king/home
kin@cse.cuhk.edu.hk



Pleas send Proposals to either the Series Editor or Directly to :

LEONG Li-Ming
Editor, CRC Press
240 Macpherson Road, #08-01,
Pines Industrial Building,
S'pore 348574
li.ming@tandf.com.sg
Tel: (65) 67415166 x 115

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

VeriGuide Free Trial

IRWIN KING @ WEB INTELLIGENCE & SOCIAL COMPUTING LAB

Trace: » home

You are here: home

NAVIGATION

- Home
- Profile
- Research Interests & Projects

ABOUT US

- Research Group | Presentations
- News | Newsletter
- Collaborators
- Contact Us

PUBLICATIONS

1. Conference Papers 2010-Now
2. Conference Papers 2005-2009
3. Journal Articles
4. Books, Edited Books & Proceedings
5. Book Chapters
6. Conference Papers 2000-2004
7. Conference Papers 1994-1999
8. Theses
9. Patents
10. Presentations
11. Datasets

PROFESSIONAL ACTIVITIES

1. Professional Achievements
2. Awards and Recognitions
3. Grants
4. Teaching
5. Education Excellence
6. Demos & Software
   I. Finding Experts Demo
   II. MEMPM Matlab Toolbox
   III. My iApps
7. Conference Activities

Irwin King, WISC Lab

"...the truth shall set you free.", Caltech's Motto

Professor, B.Sc. (Caltech), M.Sc., Ph.D. (USC)
SMIEEE (CIS), MACM, SMINNS (VP, BoG), APNNA (VP,BoG), IrwinKing.com

Department of Computer Science and Engineering
The Chinese University of Hong Kong, Shatin, NT, Hong Kong
Phone: +(852) 3943 8398; **Fax**: +(852) 2603 5024
Email: king [at] cse [dot] cuhk [dot] edu [dot] hk

- Visiting Professor with School of Information (iSchool), UC Berkeley (2011-2012)
- AT&T Lbas AT&T Labs Research, San Francisco (2010-2012)

- Book Series Editor, Social Media and Social Computing, Taylor and Francis (CRC Press)
- Associate Editor of ACM Transactions on Knowledge Discovery from Data (ACM TKDD)
- Associate Editor of INNS Natural Intelligence Magazine (INNS NIM)
- Associate Editor of IEEE Transactions on Neural Networks (IEEE TNN)
- Vice-President of Membership, Board Member, Board of Governors, International Neural Network Society (INNS)
- Vice-President and Board Member, Asia Pacific Neural Network Assembly (APNNA)
- Chair, Task Force on the Future Directions of Neural Networks (IEEE CIS)
- Chair, SIG and Regional Chapters Committee for Asia and the Pacific (INNS)
- Member, CIS Outstanding Chapter Award Subcommittee, (IEEE CIS)
- Member of Review Panel of the Natural Science, and Engineering of Academy of Finland
- Member of Review Panel of the Natural Sciences and Engineering Research Council of Canada (NSERC)
- Member of RGC Engineering Panel, The Hong Kong SAR Government
- Member of Joint Research Scheme (JRS) Panel under RGC, The Hong Kong SAR Government
- Principal Investigator, Chief Technologist, and Co-Founder, The VeriGuide Project, CUHK
- Member of the Engineering Faculty Board, The Chinese University of Hong Kong
- Member of the Editorial Board, Web Intelligence and Web Science (WIWS), Higher Education Press, China
- Kavli Fellow, Kavli Frontiers of Science Symposium, Kavli Foundation
- **Special Issue Guest Editor**, Twitter and Microblogging Services, ACM Transactions on Intelligent System and Technology

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

Acknowledgments

- Shouyuan Chan (Ph.D.)
- Chen Cheng (Ph.D.)
- Baichuan Li (Ph.D.)
- Guang Ling (Ph.D.)
- Haiqin Yang (Postdoc)
- Connie Yuen (Ph.D.)
- Hongyi Zhang (Ph.D.)
- Chao Zhou (Ph.D.)
- Patrick Lau
- Raymond Yeung


The Chinese University of Hong Kong

CUHK Excellence

- The only university in Hong Kong having Nobel Laureates as faculty. **Four** Distinguished Professors-at-Large.

Professor Yang Chen-Ning, Nobel Laureate in Physics

Professor Sir James A. Mirrlees, Nobel Laureate in Economic Sciences

Professor Yau Shing-Tung, Fields Medalist

Professor Andrew Yao, Turing Award Winner

- **Nine** academicians of Chinese Academy of Sciences and Chinese Academy of Engineering

Q & A