CMSC5733 Social Computing

Irwin King

The Chinese University of Hong Kong

king@cse.cuhk.edu.hk

Social Recommender Systems

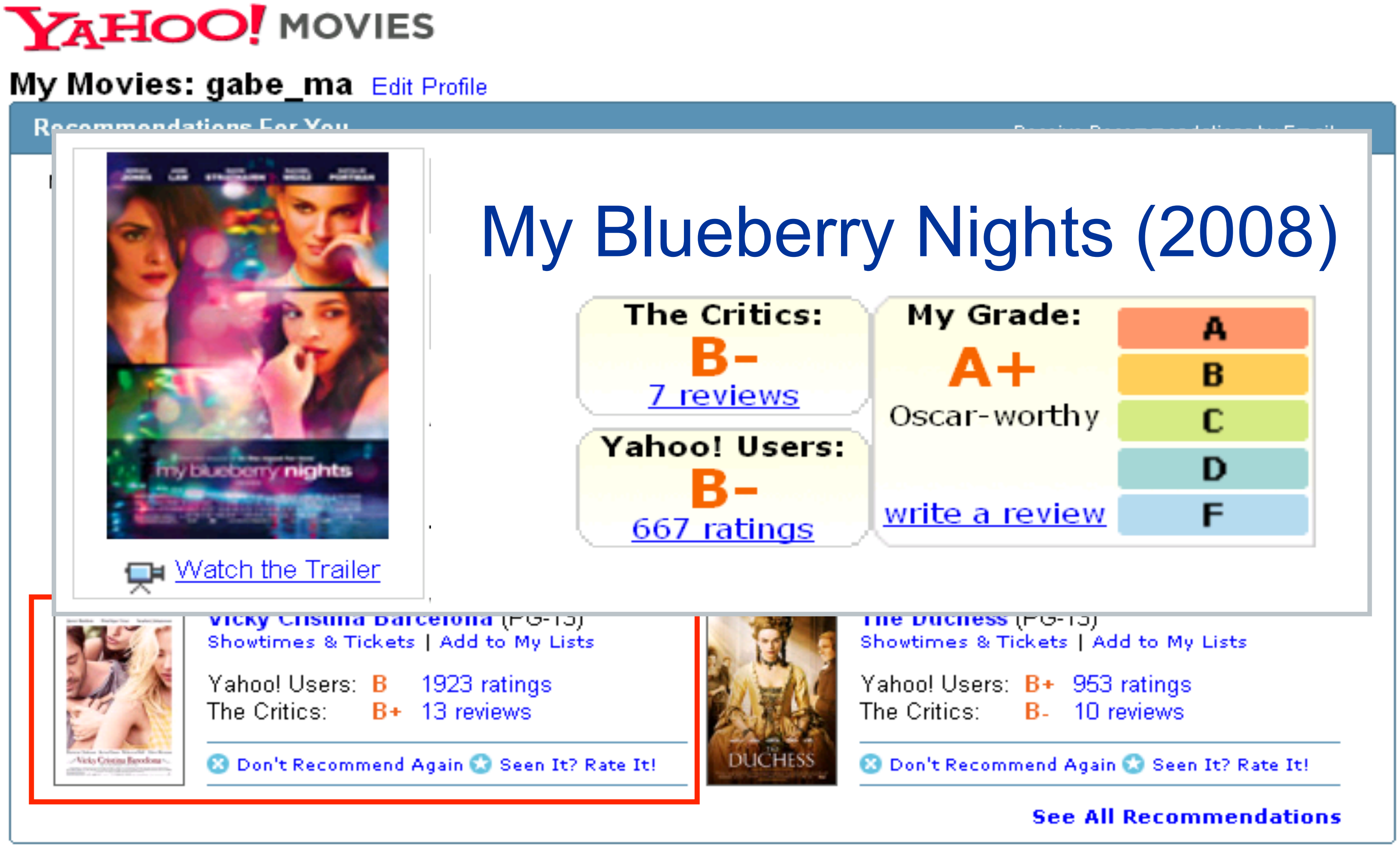
- Introduction
- Collaborative Filtering
- Trust-aware Recommender Systems
- Social-based Recommender Systems
- Web Site Recommendation

Social Recommendation Using Probabilistic Matrix Factorization

[Hao Ma, et al., CIKM2008]

Challenges

- Data sparsity problem



Challenges

My Movie Ratings

The Pursuit of Happyness (PG-13, 1 hr. 57 min.)
Buy DVD | Add to My Lists

Yahoo! Users:	B+	38992 ratings
The Critics:	B-	13 reviews

My Rating: A+

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+

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

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

Cold Mountain (R, 2 hrs. 35 min.)
Buy DVD | Add to My Lists

Yahoo! Users:	B	38986 ratings
The Critics:	B+	10 reviews

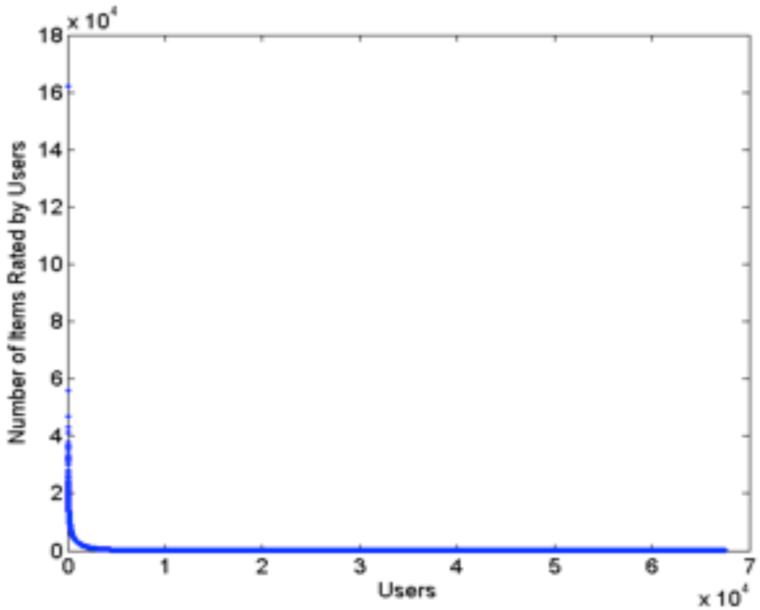
My Rating: B+

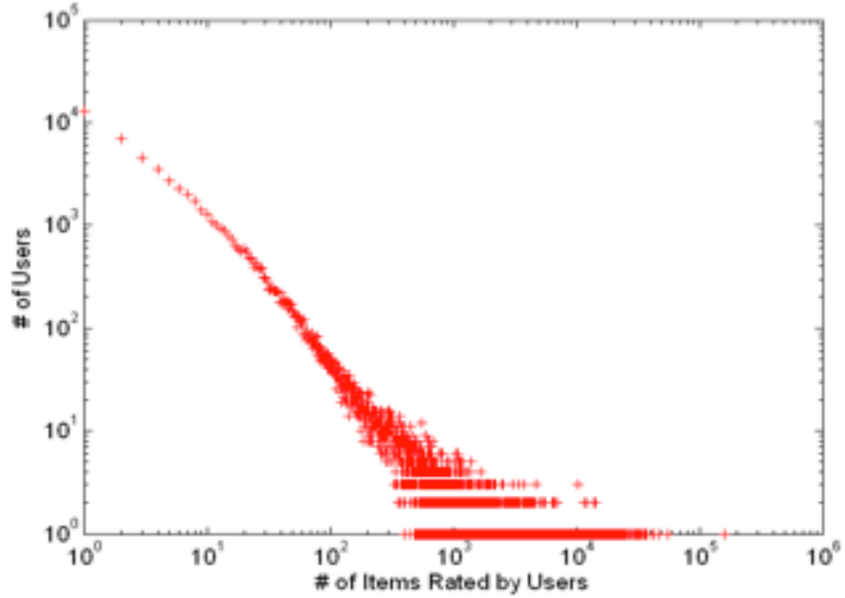
Shrek 2 (PG, 1 hr. 32 min.)
Buy DVD | Add to My Lists

Yahoo! Users:	B+	150368 ratings
The Critics:	B	15 reviews

My Rating: B

Number of Ratings per User





Extracted From Epinions.com

114,222 users, 754,987 items and 13,385,713 ratings

Challenges

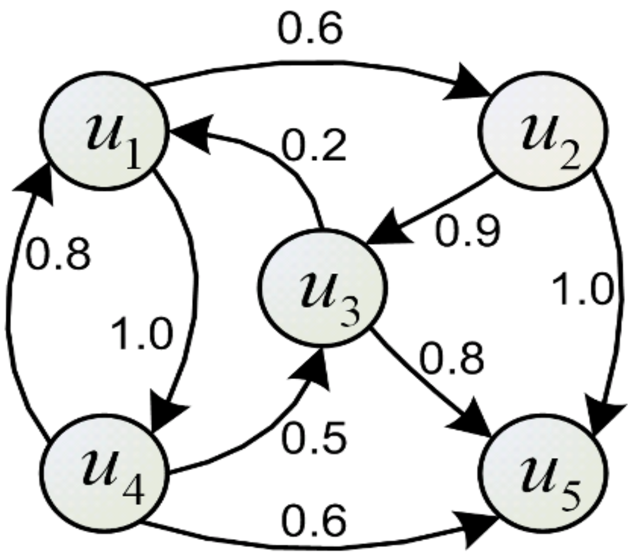
- Traditional recommender systems ignore the social connections between users

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



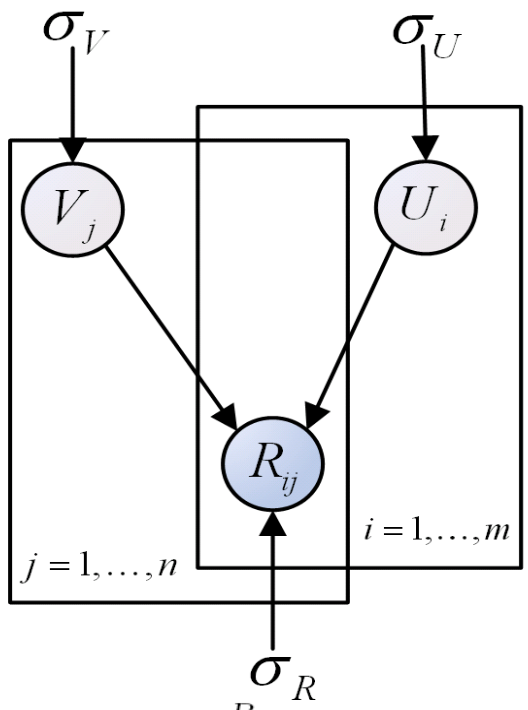
Social Trust Graph

	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



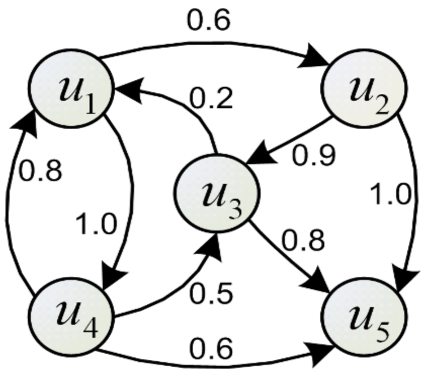
$$p(R|U, V, \sigma_R^2) = \prod_{i=1}^{m} \prod_{j=1}^{n} \left[\mathcal{N}\left(R_{ij} | g(U_i^T V_j), \sigma_R^2 \right) \right]^{I_{ij}^R}$$

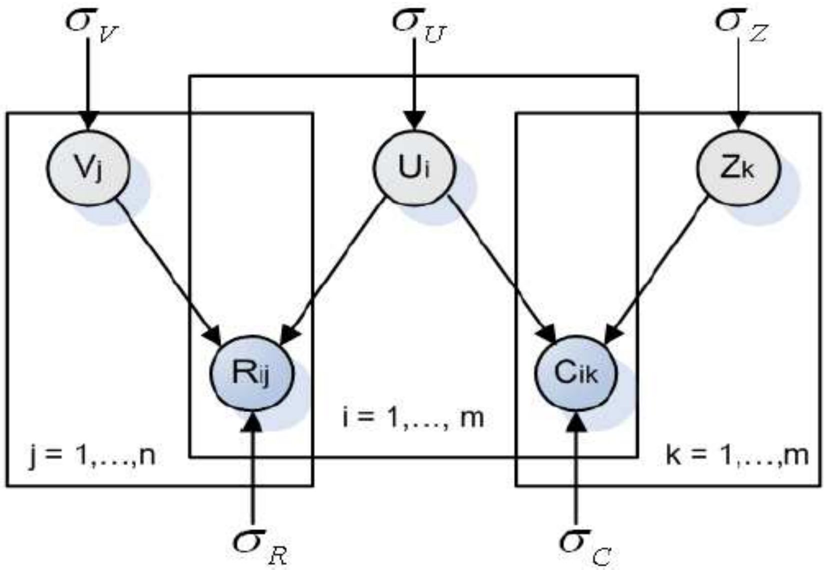
$$p(U|\sigma_U^2) = \prod_{i=1}^{m} \mathcal{N}(U_i | 0, \sigma_U^2 \mathbf{I}) \qquad p(V|\sigma_V^2) = \prod_{j=1}^{n} \mathcal{N}(V_j | 0, \sigma_V^2 \mathbf{I})$$

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

SoRec

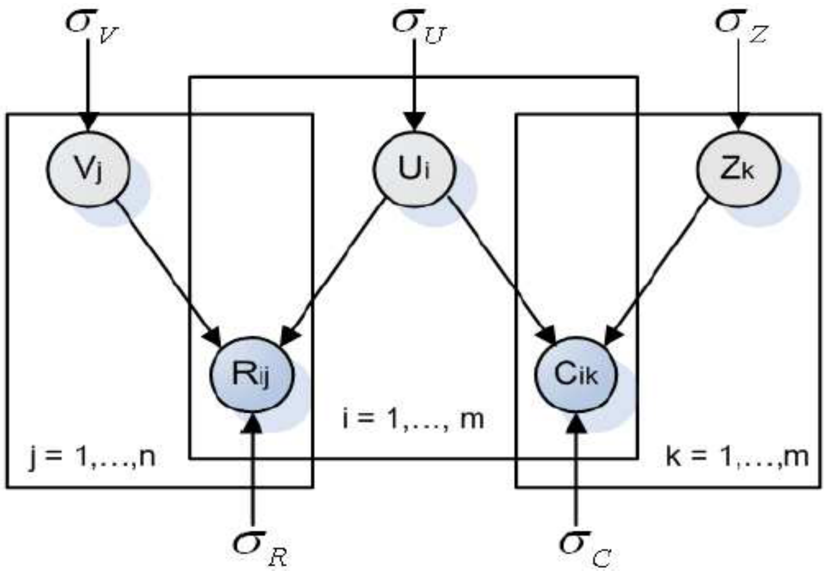
	v_1	v_2	v_3	v_4	v_5	v_6
u_1		5	2		3	
u_2	4			3		4
u_3			2			2
u_4	5			3		
u_5		5	5			3





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

Experimental Analysis

- Table: MAE comparison with other approaches (A smaller MAE value means a better performance)

Training Data	Dimensionality = 5				Dimensionality = 10			
	MMMF	PMF	CPMF	SoRec	MMMF	PMF	CPMF	SoRec
99%	1.0008	0.9971	0.9842	**0.9018**	0.9916	0.9885	0.9746	**0.8932**
80%	1.0371	1.0277	0.9998	**0.9321**	1.0275	1.0182	0.9923	**0.9240**
50%	1.1147	1.0972	1.0747	**0.9838**	1.1012	1.0857	1.0632	**0.9751**
20%	1.2532	1.2397	1.1981	**1.1069**	1.2413	1.2276	1.1864	**1.0944**

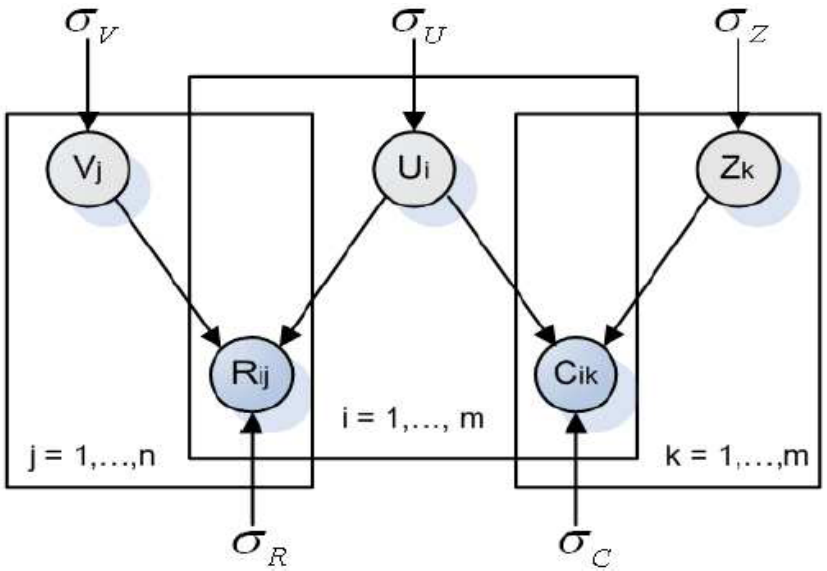
MMMF:
J.D.M Rennie and N. Srebro
(ICML'05)

PMF & CPMF:
R. Salakhutdinov and A. Mnih
(NIPS'08)

Epinions: 40,163 users who rated 139,529 items with totally 664,824 ratings

Disadvantages of SoRec

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

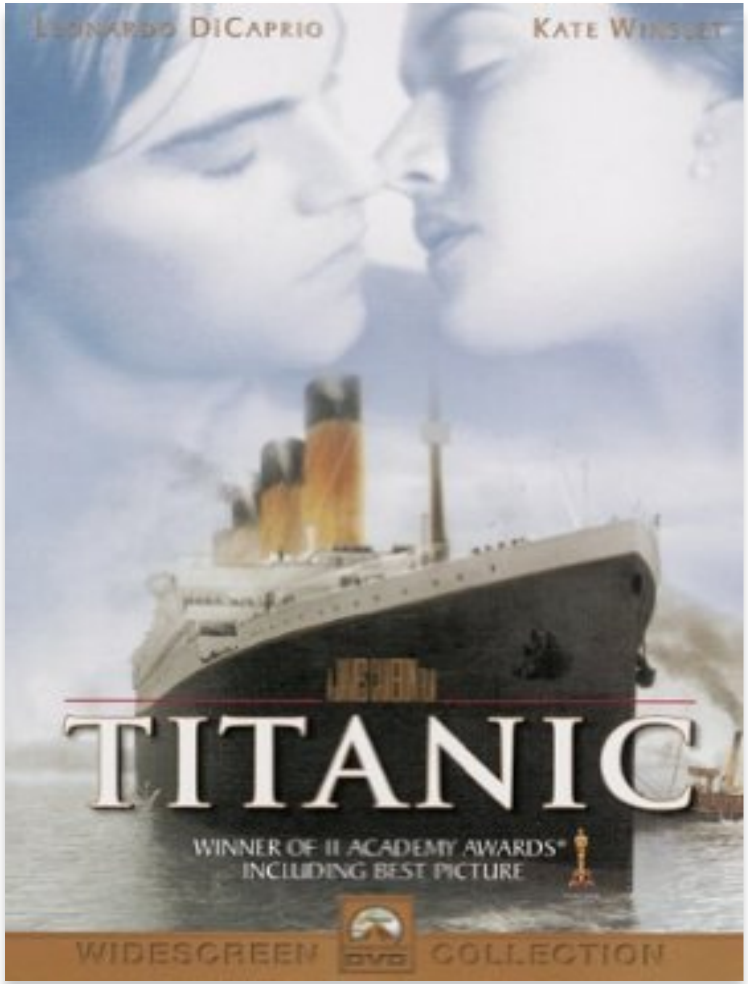


SoRec

Learning to Recommend with Social Trust Ensemble

[Hao Ma, et al., SIGIR2009]

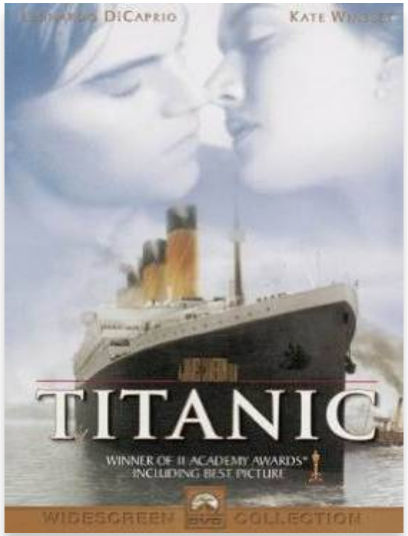
1st Motivation



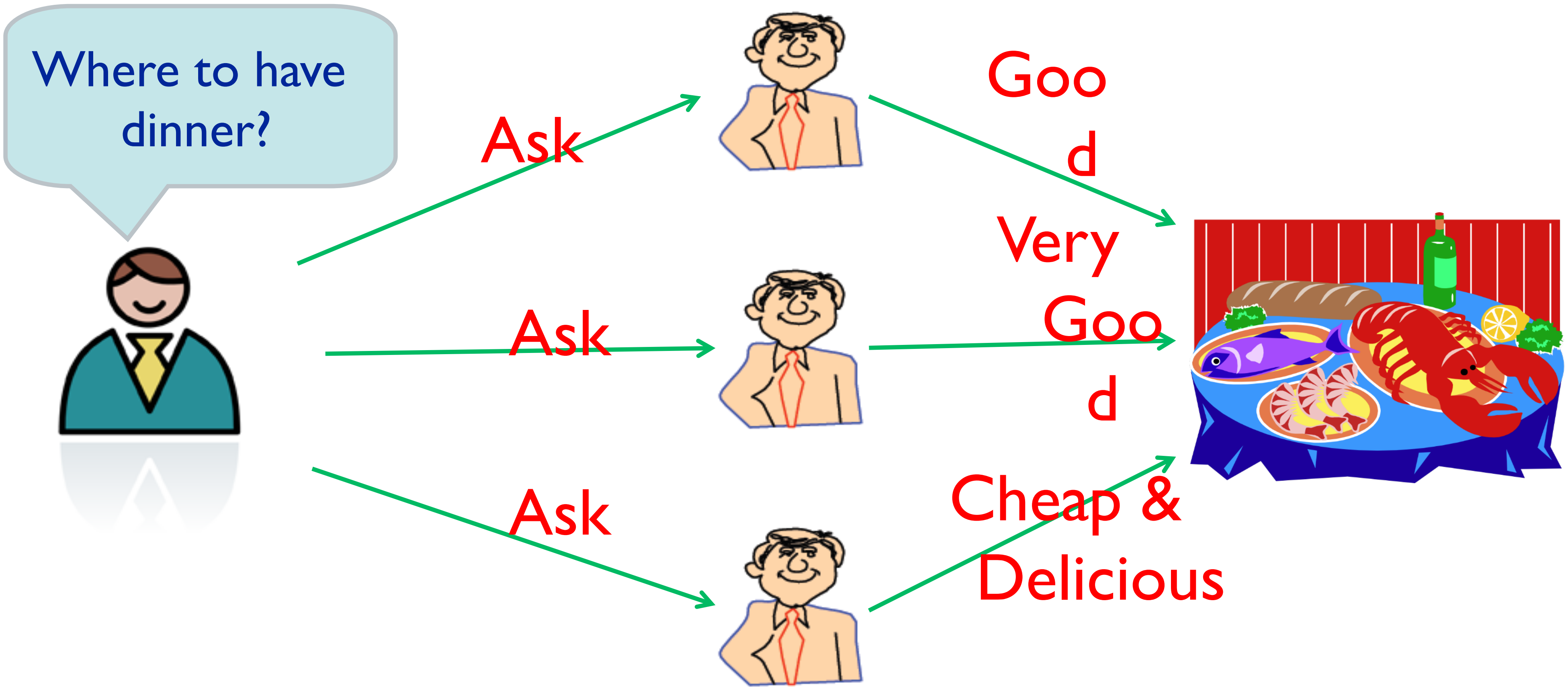
1st Motivation

1st Motivation

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

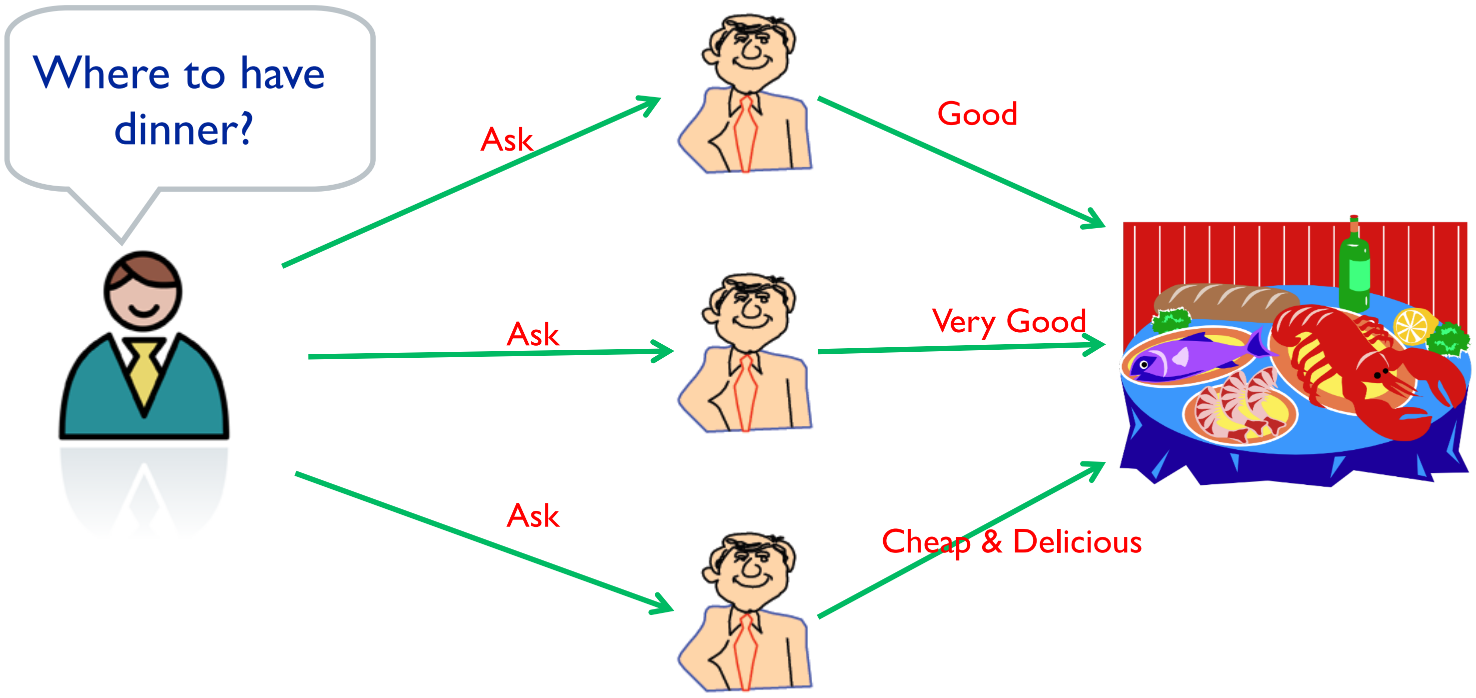


2nd Motivation



2nd Motivation

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



Motivations

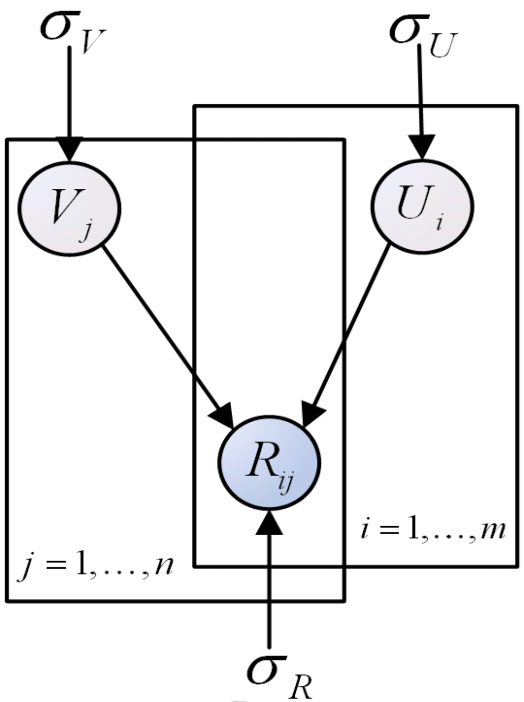
- Users have their own characteristics, and they have different tastes on different items, such as movies, books, music, articles, food, etc.

- 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



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

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

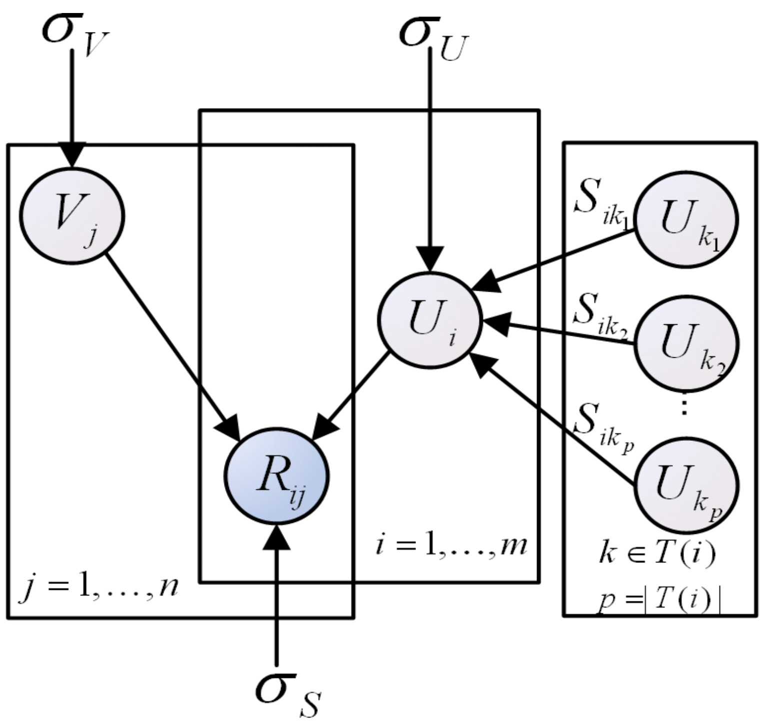
[R. Salakhutdinov, et al., NIPS2008]

Recommendations by Trusted Friends

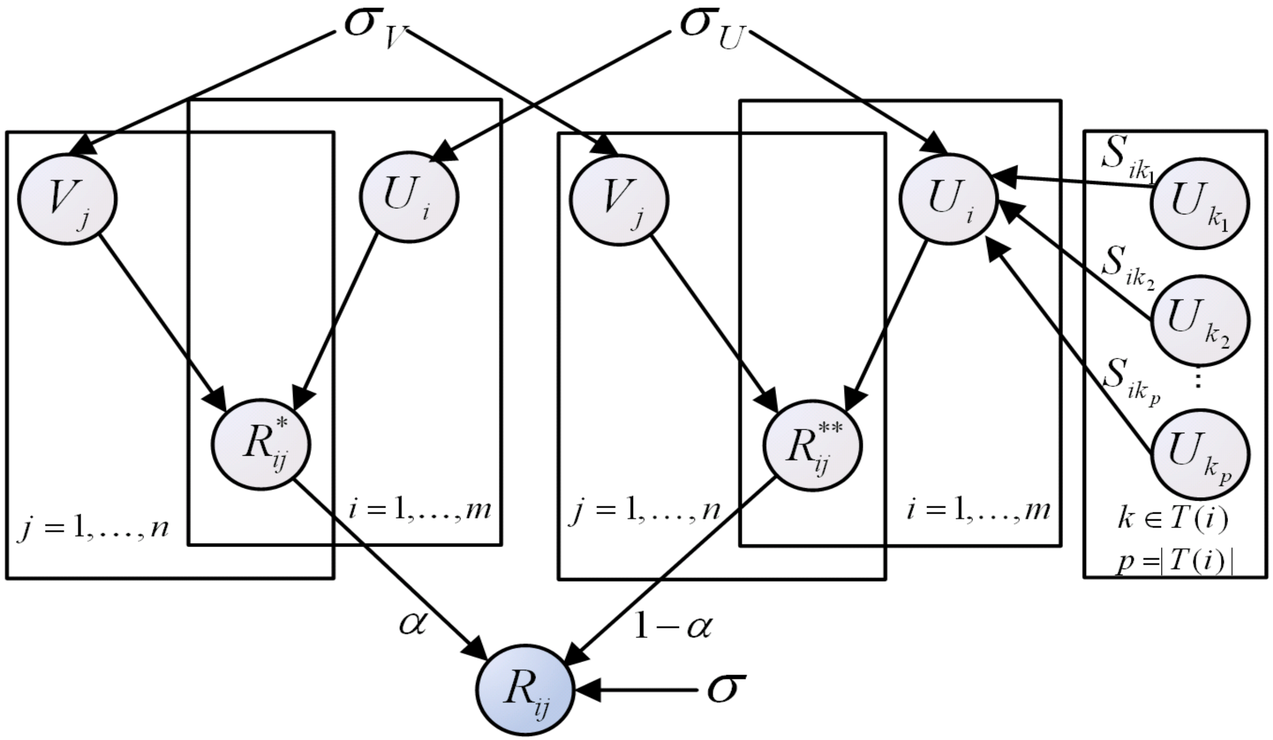
$$\widehat{R}_{ik} = \frac{\sum_{j \in \mathcal{T}(i)} R_{jk} S_{ij}}{|\mathcal{T}(i)|}$$

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

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



Recommendation with Social Trust Ensemble



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

Recommendation with Social Trust Ensemble

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

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

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

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 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**

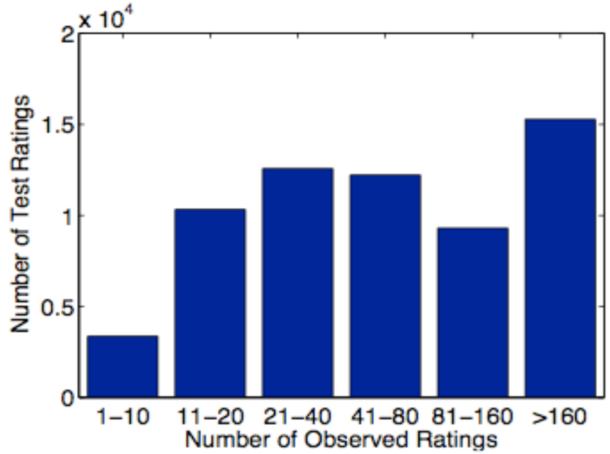
NMF --- D. D. Lee and H. S. Seung (Nature 1999)
PMF --- R. Salakhutdinov and A. Mnih (NIPS 2008)
SoRec --- H. Ma, H. Yang, M. R. Lyu and I. King (CIKM 2008)
Trust, RSTE --- H. Ma, I. King and M. R. Lyu (SIGIR 2009)

Performance on Different Users

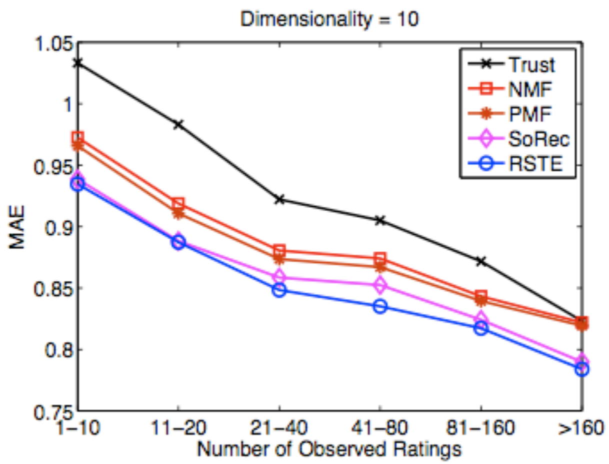
- Group all the users based on the number of observed ratings in the training data

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

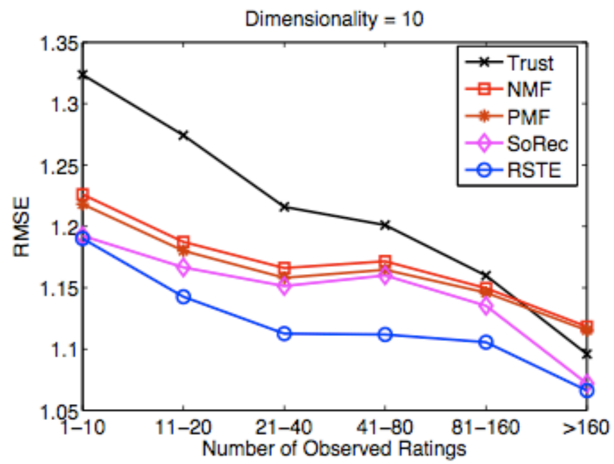
Performance on Different Users



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

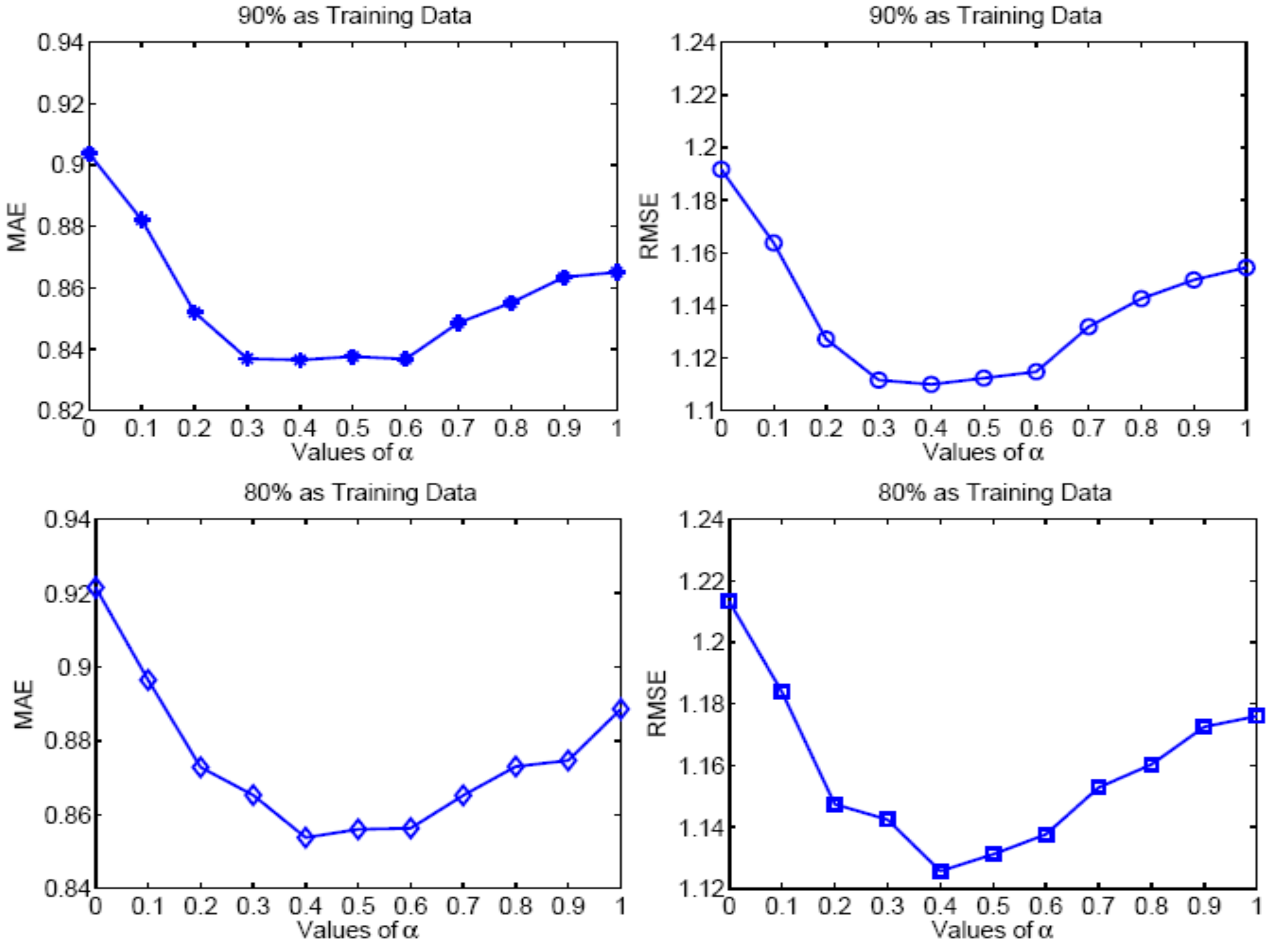


(b) MAE Comparison on Different User Rating Scales (90% as Training Data)



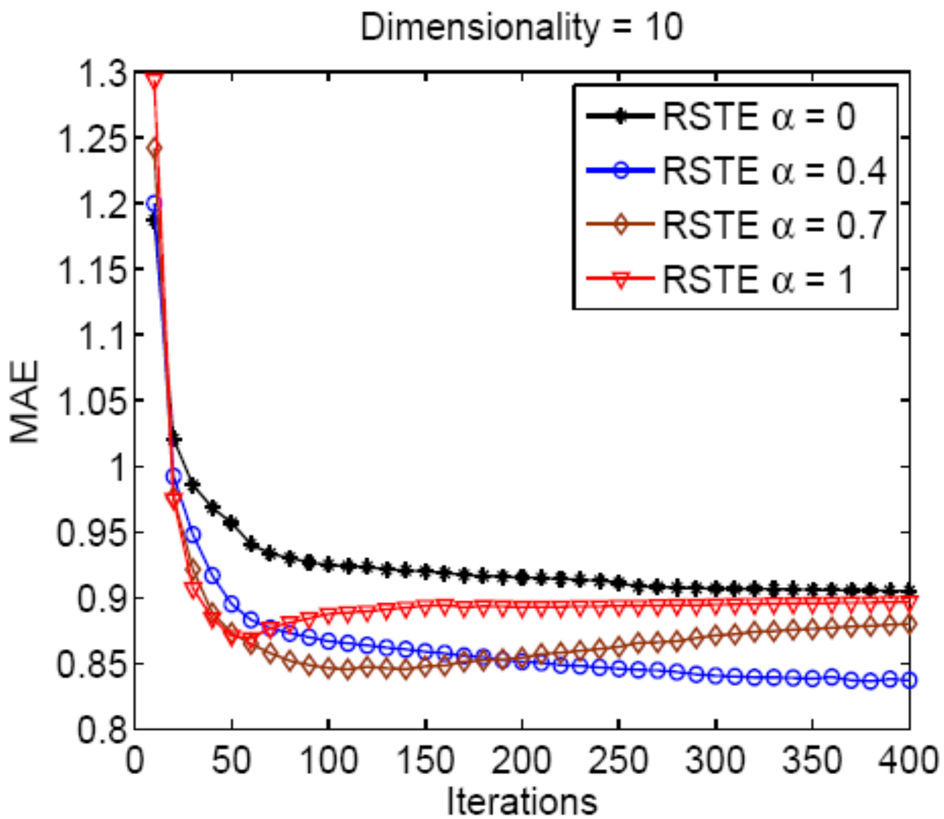
(c) RMSE Comparison on Different User Rating Scales (90% as Training Data)

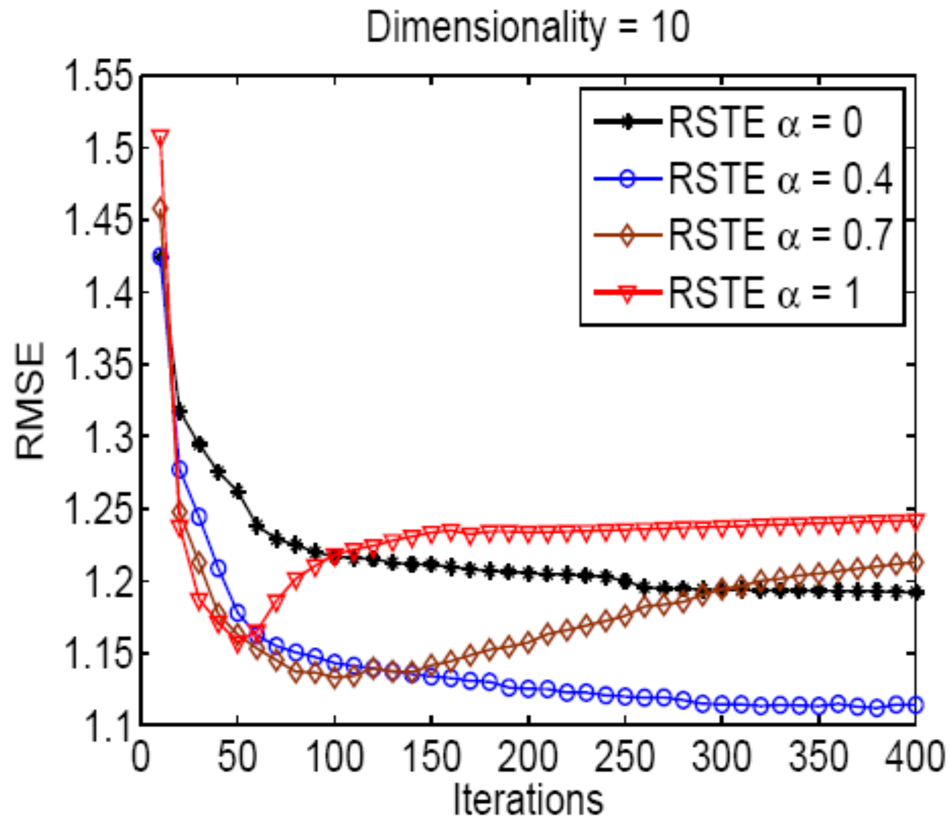
Impact of Parameter Alpha



Impact of Parameter α (Dimensionality = 10)

MAE and RMSE Changes with Iterations

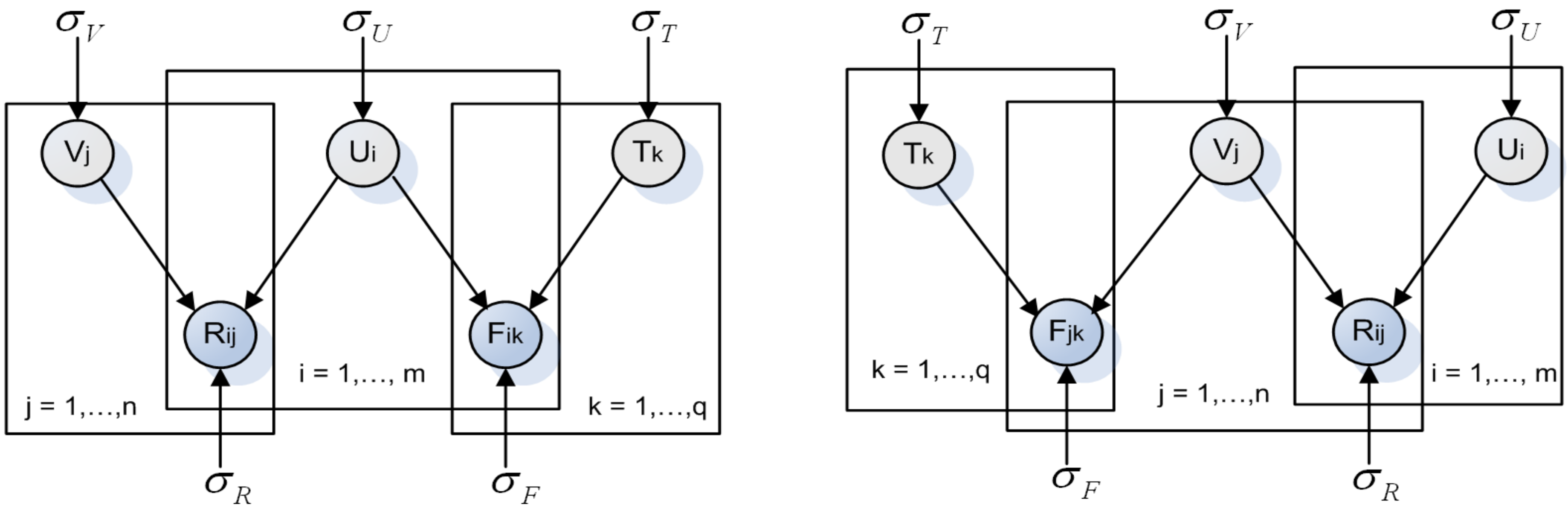




90% as Training Data

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

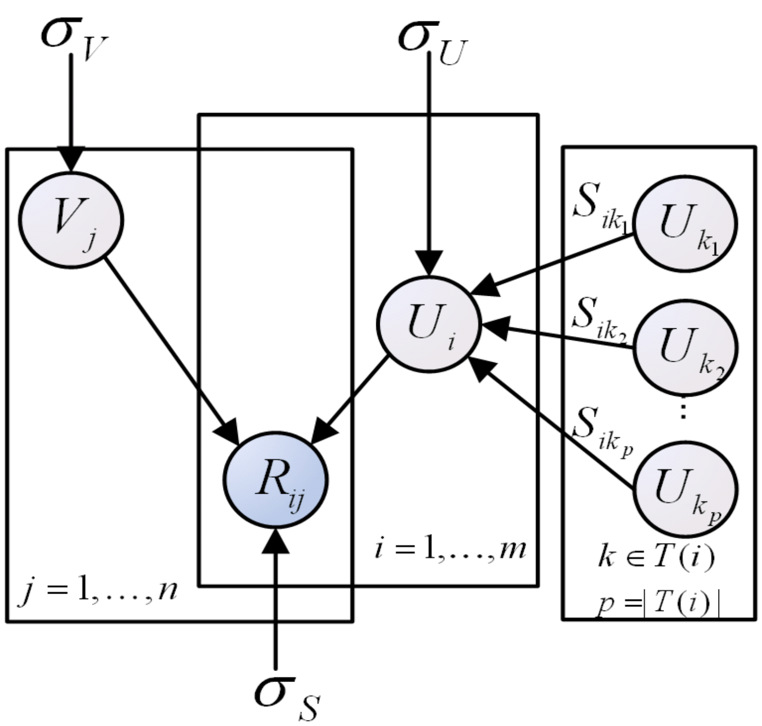
- RMSE

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

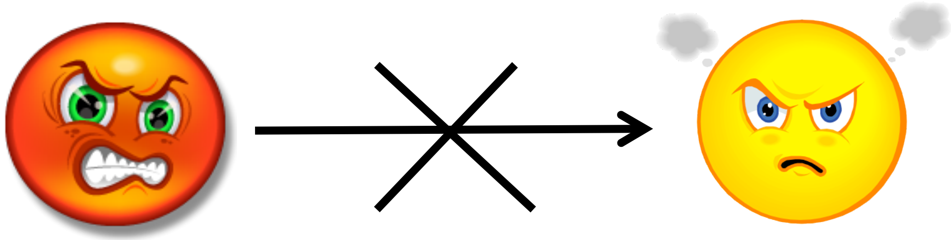


- The trusted friends are actually the explicit neighbors
- We can easily apply this method to include implicit neighbors
- Using PCC to calculate similar users for every user

What We Cannot Model Using SoRec and RSTE?

- Propagation of trust

- Distrust



Social Recommender Systems

- Introduction
- Collaborative Filtering
- Trust-aware Recommender Systems
- Social-based Recommender Systems
- Web Site Recommendation

Recommend with Social Distrust

[Hao Ma, et al., RecSys2009]

Trust vs. Social

- Trust-aware
  - Trust network: unilateral relations
  - Trust relations can be treated as "similar" relations
  - Few datasets available on the Web
- Social-based
  - Social friend network: mutual relations
  - Friends are very diverse, and may have different tastes
  - Lots of Web sites have social network implementation

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

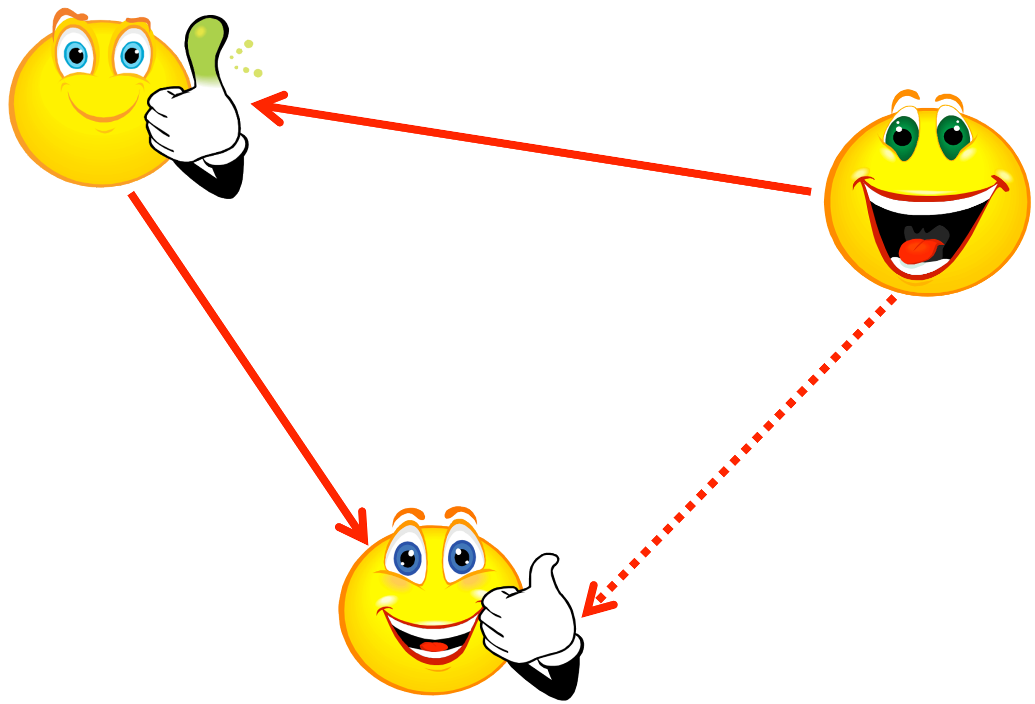
- Users' trust relations can be interpreted as the "similar" relations
  - On the web, user U_i trusts user U_t indicates that user U_i agrees with most of the opinions issued by user U_t.

Trust

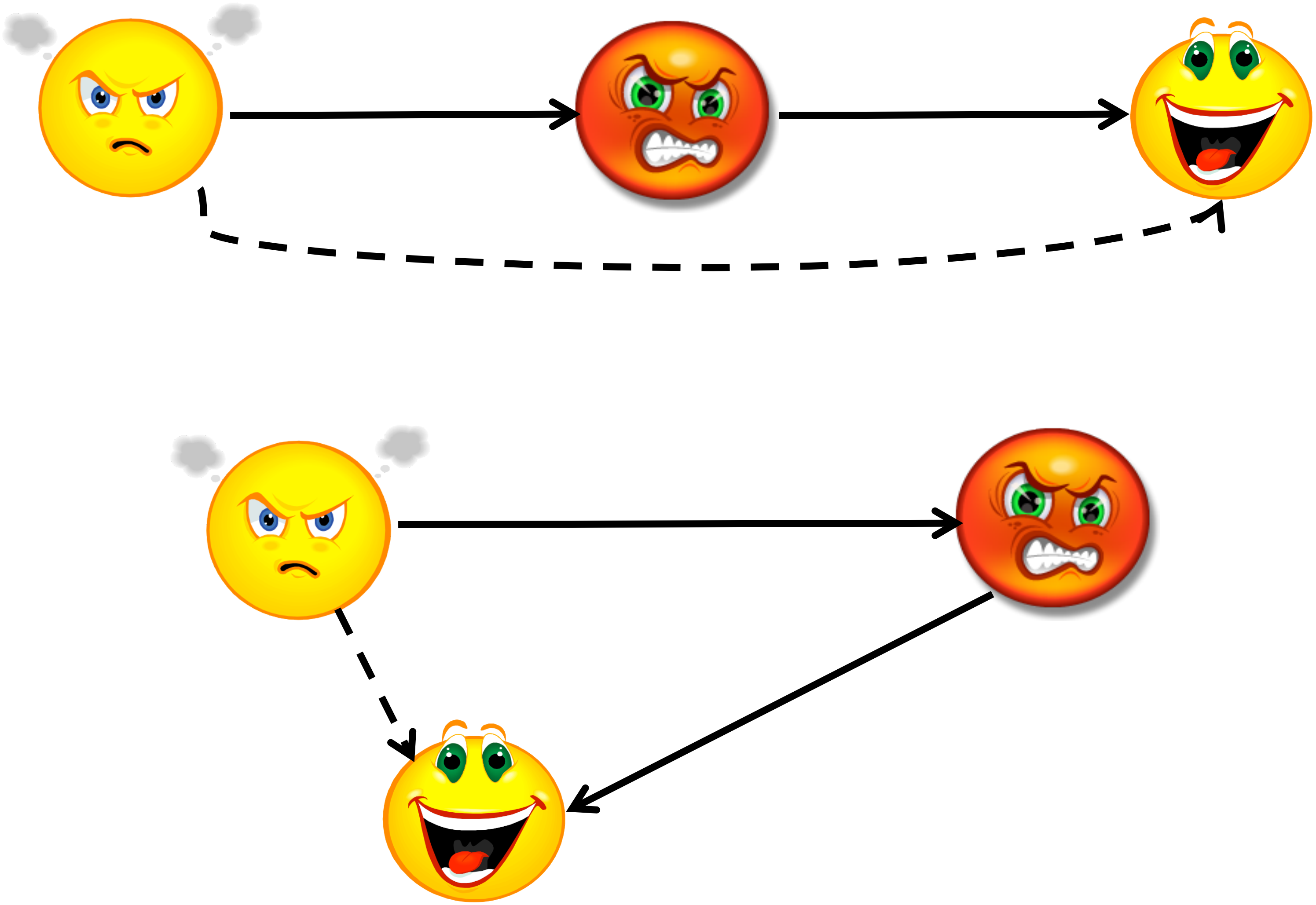
$$\min_{U} \frac{1}{2}\sum_{i=1}^{m}\sum_{t\in\mathcal{T}^{+}(i)} S_{it}^{\mathcal{T}} \|U_i - U_t\|_F^2$$

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

Trust Propagation



Distrust Propagation?



Experiments

- Dataset - Epinions
- 131,580 users, 755,137 items, 13,430,209 ratings
- 717,129 trust relations, 123,670 distrust relations

Data Statistics

Table 1: Statistics of User-Item Rating Matrix of Epinions

Statistics	User	Item
Min. Num. of Ratings	1	1
Max. Num. of Ratings	162169	1179
Avg. Num. of Ratings	102.07	17.79

Table 2: Statistics of Trust Network of Epinions

Statistics	Trust per User	Be Trusted per User
Max. Num.	2070	3338
Avg. Num.	5.45	5.45

Table 3: Statistics of Distrust Network of Epinions

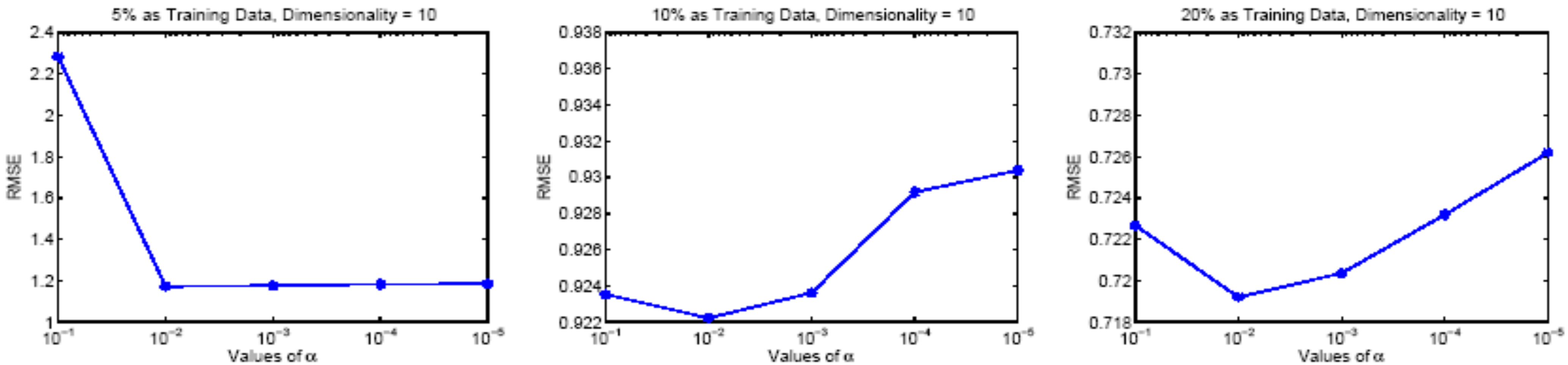
Statistics	Distrust per User	Be Distrusted per User
Max. Num.	1562	540
Avg. Num.	0.94	0.94

Experiments

RMSE

Dataset	Traning Data	Dimensionality	PMF	SoRec	RWD	RWT
Epinions	5%	5D	1.228	1.199	1.186	**1.177**
		10D	1.214	1.198	1.185	**1.176**
	10%	5D	0.990	0.944	0.932	**0.924**
		10D	0.977	0.941	0.931	**0.923**
	20%	5D	0.819	0.788	0.723	**0.721**
		10D	0.818	0.787	0.723	**0.720**

Impact of Parameters



(a) 5% as Training Data (b) 10% as Training Data (c) 20% as Training Data

Figure 6: Impact of Parameter α

Alpha = 0.01 will get the best performance!

Parameter beta basically shares the same trend!

Social Recommender Systems

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

Comparison

- Trust-aware Recommender systems
  - Trust network
  - Trust relations can be treated as “similar” relations
  - Few dataset available on the web
- Social-based Recommender Systems
  - Social friend network, mutual relations
  - Friends are very divers, and may have different tastes
  - Lots of web sites have social network implementation

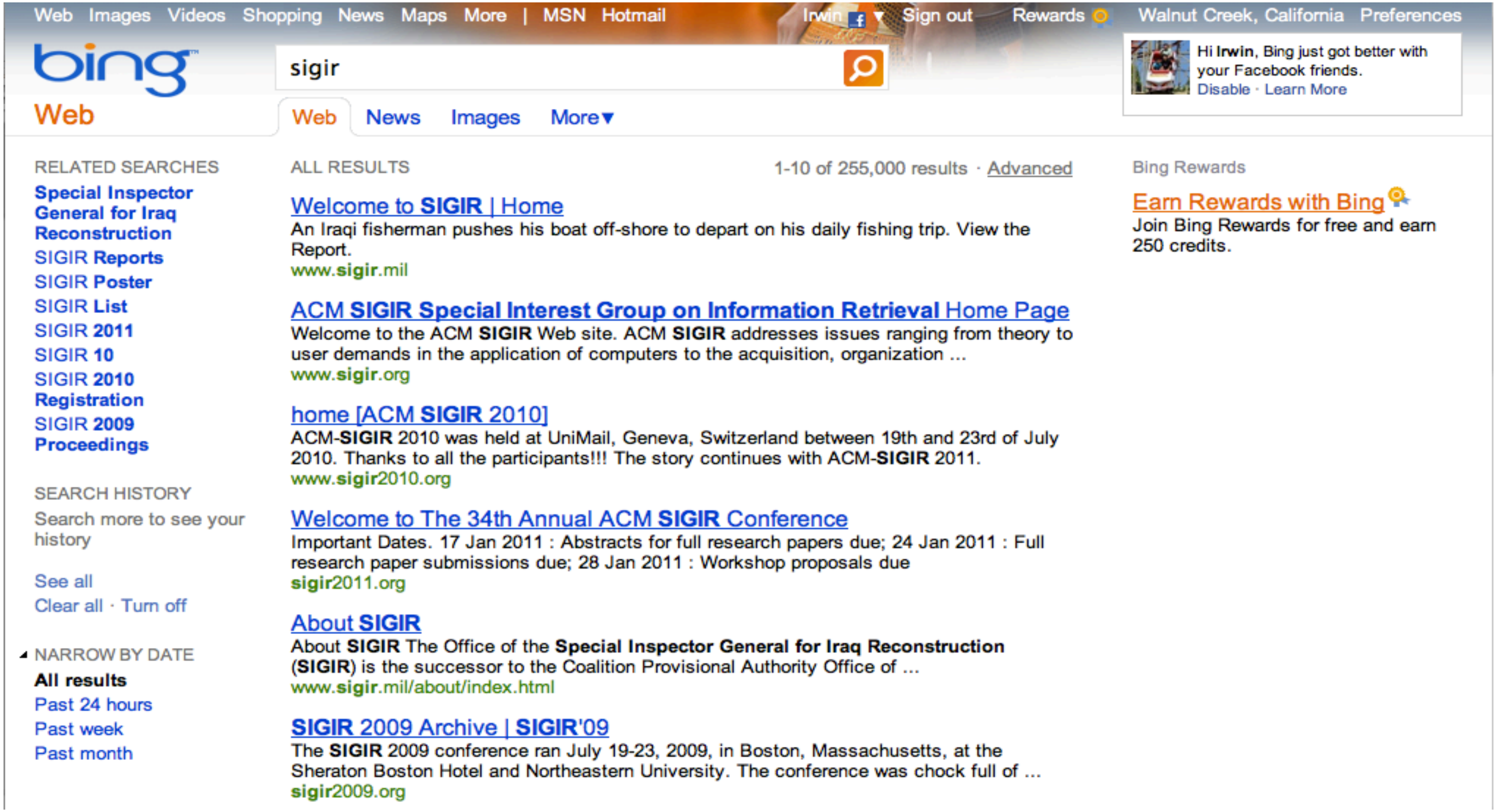
Social Recommender Systems

- Introduction
- Collaborative Filtering
- Trust-aware Recommender Systems
- Social-based Recommender Systems
- Web Site Recommendation

Web Site Recommendation

[Ma et al., SIGIR 2011]

Traditional Search Paradigm



“Search” to “Discovery”

Challenges

- 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

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

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

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

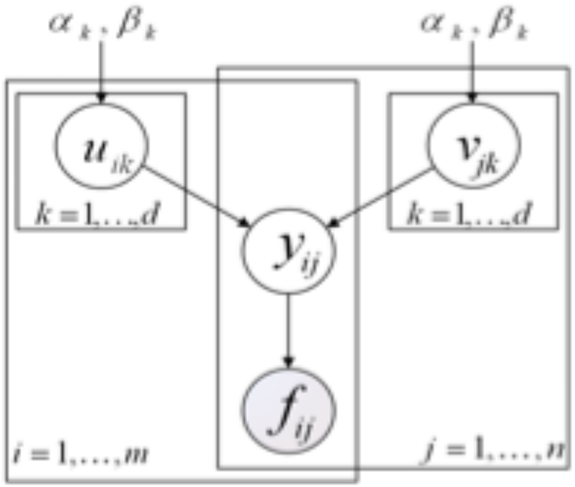
$$p(V|\sigma_V^2) = \prod_{j=1}^{n} \mathcal{N}(V_j | 0, \sigma_V^2 \mathbf{I})$$

Probabilistic Factor Model

- GaP [Canny, SIGIR 2004]
  - Linear topic model

- NMF
  - No Gamma distribution on X

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

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

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

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

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

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

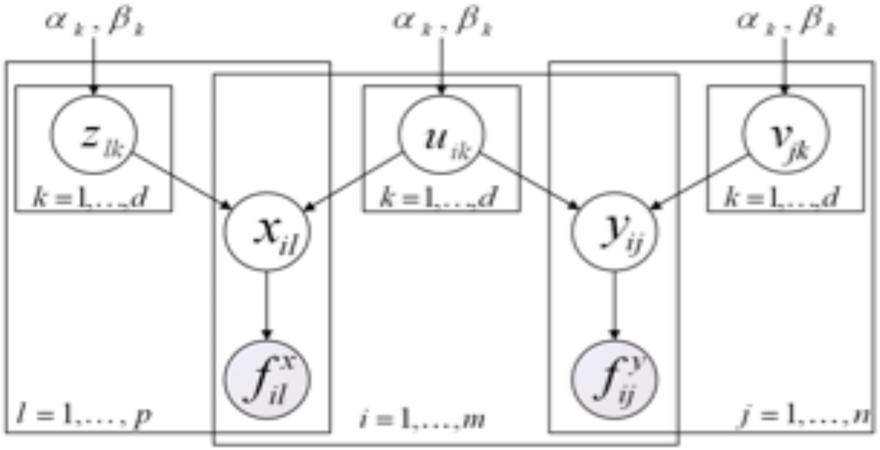
Probabilistic Factor Model

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

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

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

Collective Probabilistic Factor Model



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

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

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

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

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

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

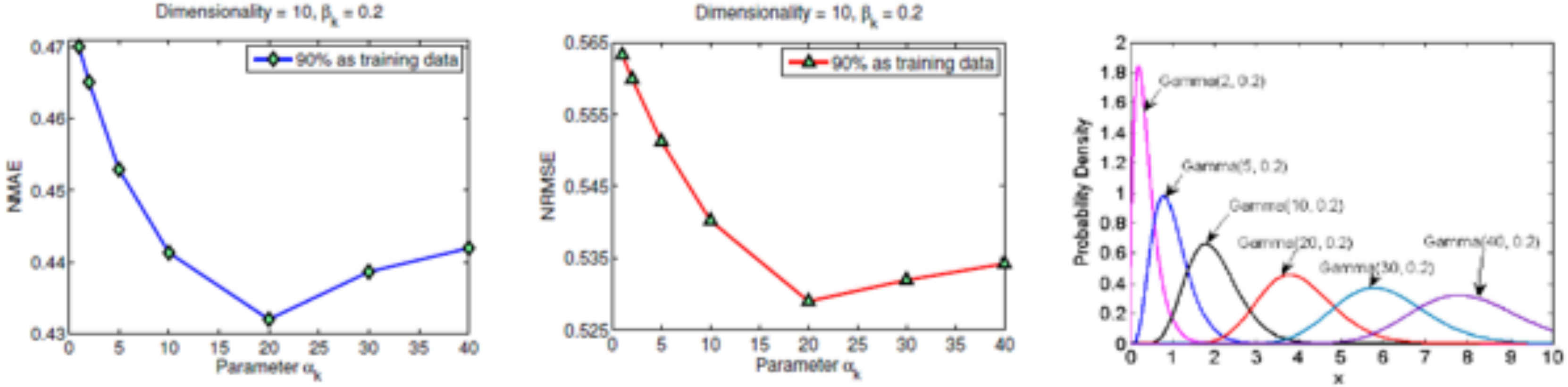
Table 3: Performance Comparison (Dimensionality = 10)

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%		
	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%		

Table 4: Performance Comparison (Dimensionality = 20)

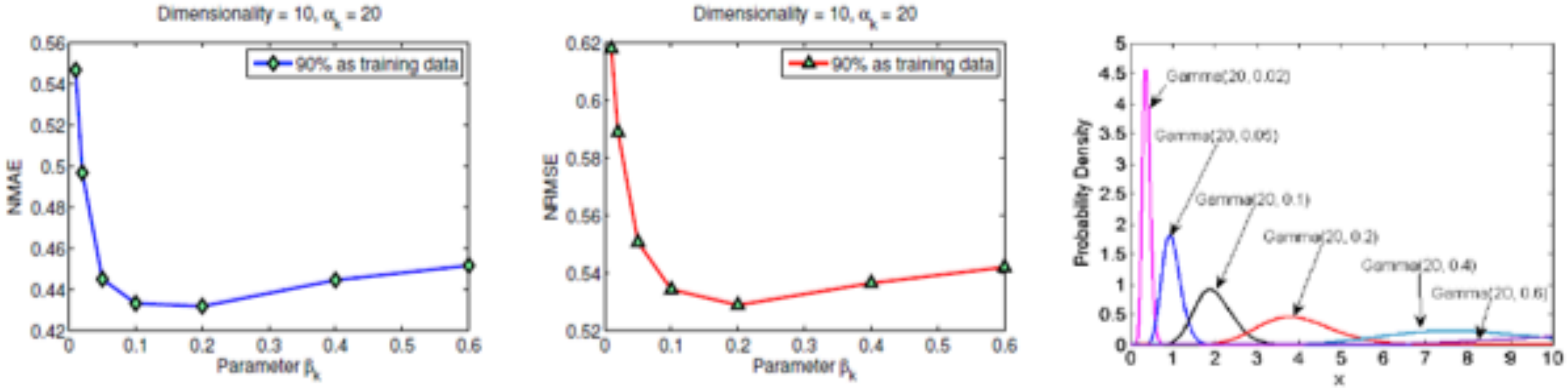
Training Data	Metrics	UserMean	SiteMean	SVD	PMF	NMF	GaP	PFM	CPFM
90%	NMAE	2.246	1.094	0.469	0.460	0.449	0.426	0.413	**0.409**
	Improve	81.79%	62.61%	12.79%	11.09%	8.91%	3.99%		
	NRMSE	3.522	2.171	0.568	0.556	0.542	0.521	0.503	**0.496**
	Improve	85.92%	77.15%	12.68%	10.79%	8.49%	4.80%		
80%	NMAE	2.252	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%		
	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



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

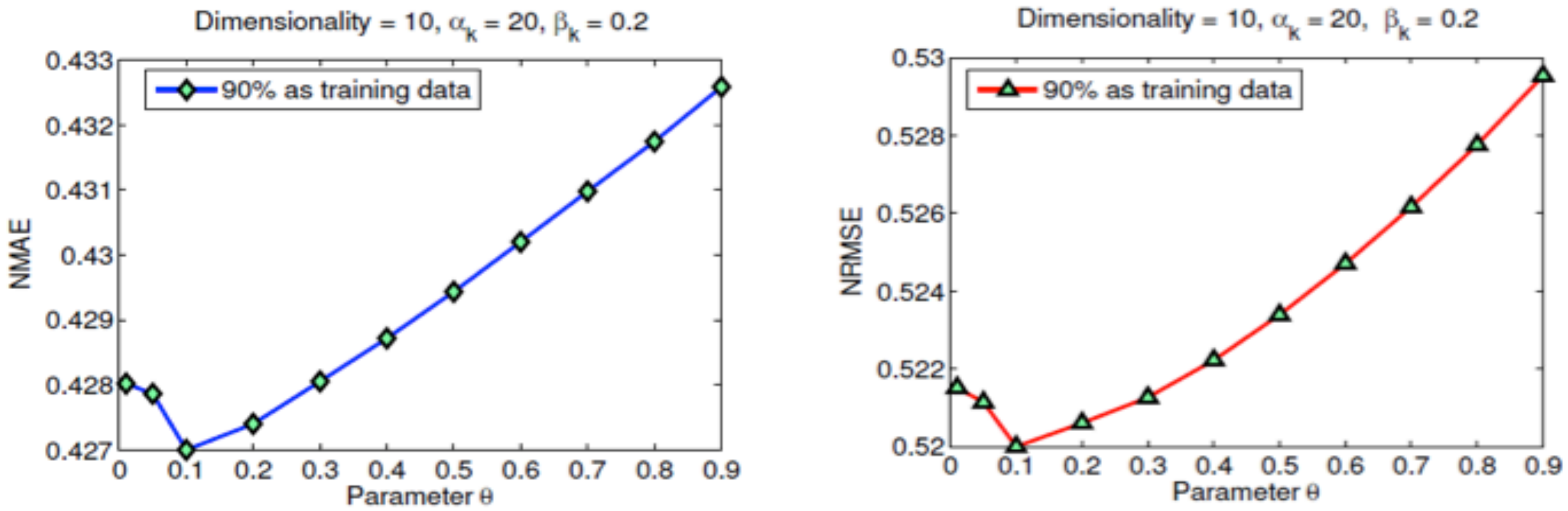
Figure 6: Impact of Parameter α_k in PFM



(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

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 patters, temporal dynamics, location information, etc. provides better models for social recommendations

References

- J. Basilico and T. Hofmann. Unifying collaborative and content-based filtering. In ICML, 2004.
- J. S. Breese, D. Heckerman, and C. M. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In UAI, pages 43–52, 1998.
- M. Deshpande and G. Karypis. Item-based top-N recommendation algorithms. ACM Trans. Inf. Syst., 22(1):143–177, 2004.
- J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl. An algorithmic framework for performing collaborative filtering. In SIGIR, pages 230–237. ACM, 1999.
- J. L. Herlocker, J. A. Konstan, and J. Riedl. An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms. Inf. Retr., 5(4):287–310, 2002.
- G. Linden, B. Smith, and J. York. Industry report: Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Distributed Systems Online, 4(1), 2003.

References

- H. Ma, I. King, and M. R. Lyu. Effective missing data prediction for collaborative filtering. In SIGIR, pages 39–46, 2007.
- H. Ma, H. Yang, M. R. Lyu, and I. King. SoRec: social recommendation using probabilistic matrix factorization. In CIKM, pages 931–940, 2008.
- H. Ma, I. King, and M. R. Lyu. Learning to recommend with social trust ensemble. In SIGIR, pages 203-210, 2009.
- H. Ma, M. R. Lyu, and I. King. Learning to recommend with trust and distrust relationships. In RecSys, pages 189-196, 2009.
- B. M. Sarwar, G. Karypis, J. A. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In WWW, pages 285–295, 2001.
- R. Salakhutdinov, and A. Mnih. Probabilistic Matrix Factorization. In NIPS, 2007.
- D. D. Lee, and H. S. Seung. Algorithms for Non-negative Matrix Factorization. In NIPS, pages 556-562, 2000.