Learning to Recommend with Social Trust Ensemble(RSTE)

2020-30977 Sangwoo Moon

SEOUL NATIONAL UNIV.
VISION & LEARNING

Contents

- Challenges in recommender system
- Social trust graph?
- Related works
- RSTE
- Empirical analysis
- Conclusions, future work and limitation

Contents

- **Challenges in recommender system**
- Social trust graph?
- Related works
- RSTE
- Empirical analysis
- Conclusions, future work and limitation

Challenge in Recommender System

- **Data Sparsity Problem.**
  - The density of the available ratings in commercial recommender system is often less than 1%. [20]

- **Traditional Recommender Systems ignore the social connections.**
  - In the real world, our favors can easily be **affected by the friends we trust.**
  - E.g., Rating items which recommended by girl/boy friend.
    $\rightarrow$ **Use Social Trust Graph**

[20] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl. Item-based collaborative filtering recommendation algorithms. In Proc. of WWW '01, pages 285–295, New York, NY, USA, 2001. ACM.

Contents

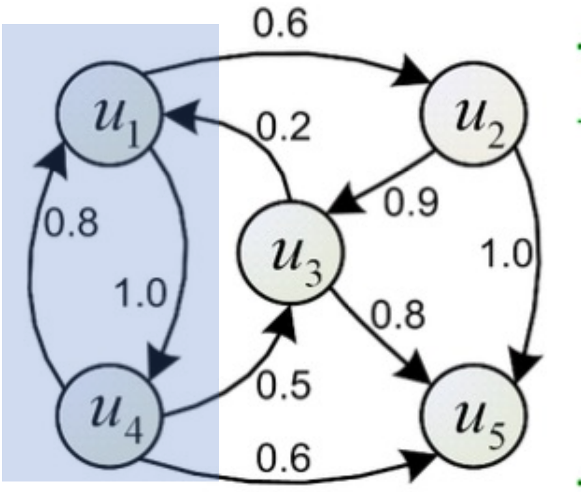
- Challenges in recommender system
- **Social trust graph?**
- Related works
- RSTE
- Empirical analysis
- Conclusions, future work and limitation

Social Trust Graph?

- Social Trust Graph
  - Node: User / Vertex: Weight of trust
  - Asymmetric directed weighted graph.

Who	Whom	Weight
1	2	0.6
1	4	1.0
2	3	0.9
2	5	1.0
…	…	…

	1	2	3	4	5
1	0	0.6	0	1.0	0
2	0	0	0.9	0	1.0
3	0.2	0	0	0	0.8
4	0.8	0	0.5	0	0.6
5	0	0	0	0	0



(a) Social Trust Graph

Social Trust Graph?

- Use dataset including Social relationship.
  (i.e., Who **trusts** whom)
- Dataset examples [1']

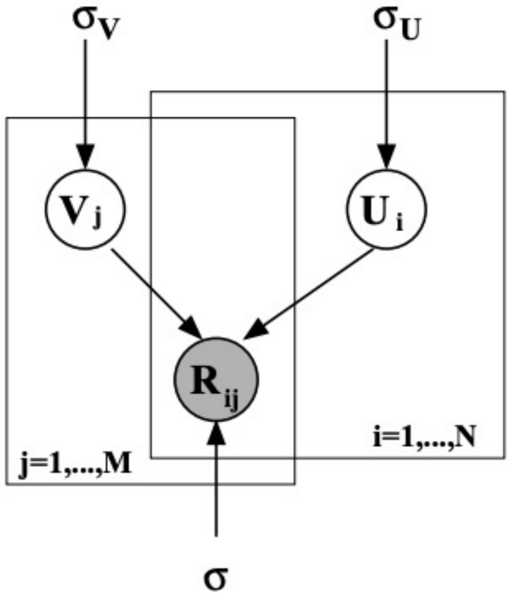
Statistics	FilmTrust	CiaoDVD	Epinions
users #	1,508	17,615	40,163
items #	2,071	16,121	139,738
ratings #	35,497	72,665	664,824
density	1.14%	0.03%	0.01%
rating range	[0.5, 4]	[1, 5]	[1, 5]
trusts #	1,853	111,781	487,183
trust density	0.42%	0.23%	0.029%

[1'] Hao Ma, Dengyong Zhou, Chao Liu, Michael R. Lyu, and Irwin King. 2011. Recommender systems with social regularization. In Proceedings of the fourth ACM international conference on Web search and data mining (WSDM '11). Association for Computing Machinery, New York, NY, USA, 287–296

Contents

[Related work] Matrix Factorization methods

- Probabilistic matrix factorization [19]
  - Define the conditional distribution over the observed rating



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

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

$$E = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M} I_{ij}\left(R_{ij} - U_i^T V_j\right)^2 + \frac{\lambda_U}{2}\sum_{i=1}^{N} \parallel U_i \parallel_{Fro}^2 + \frac{\lambda_V}{2}\sum_{j=1}^{M} \parallel V_j \parallel_{Fro}^2,$$

[19] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In Advances in Neural Information Processing Systems, volume 20, 2008.

[Related work] Trust-based Rec. Systems

- [14][15]
  - Replaces the similarity finding process with the use of a trust metric.

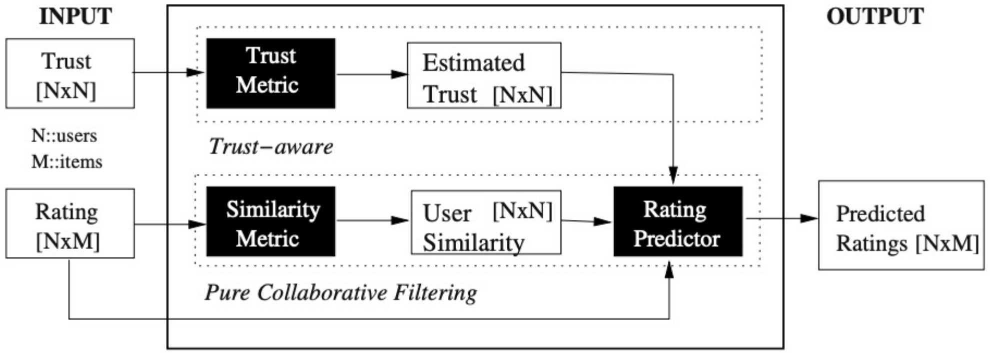


Fig. 2. Trust-Aware Recommender Systems Architecture.

  - (-) Relationship between the trust network and the user-item matrix has not been studied systematically.
  - (-) Not scalable to very large datasets, since they may need to calculate the pairwise user similarities and pairwise user trust scores.

[14] P. Massa and P. Avesani. Trust-aware collaborative filtering for recommender systems. In Proc. of CoopIS/DOA/ODBASE, pages 492–508, 2004.
[15] P. Massa and P. Avesani. Trust-aware recommender systems. In Proc. of RecSys, pages 17–24, New York, NY, USA, 2007. ACM.

[Related work] Trust-based Rec. Systems

- SoRec [13]
  - Developed factor analysis method based on the probabilistic graphical model which <u>fuses the user-item matrix with the user's social trust networks by sharing a common latent low-dimensional user feature matrix</u>.

Sharing user feature

Social network matrix

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

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

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

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

$$+\frac{\lambda_U}{2}\|U\|_F^2+\frac{\lambda_V}{2}\|V\|_F^2+\frac{\lambda_Z}{2}\|Z\|_F^2, \qquad (9)$$

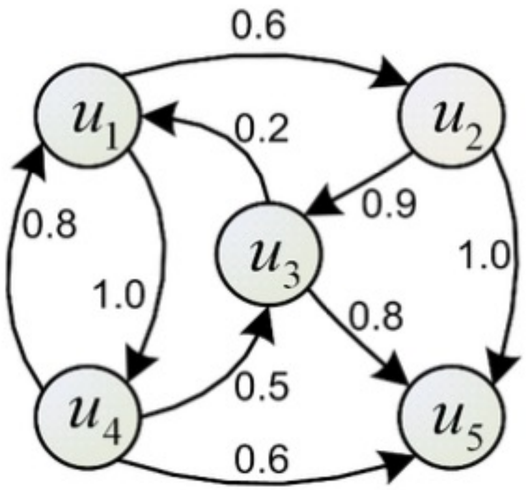
  - (-) Lack of interpretability in the model
  - (-) Recommendation qualities

[13] H. Ma, H. Yang, M. R. Lyu, and I. King. SoRec: Social recommendation using probabilistic matrix factorization. In Proc. of CIKM '08, pages 931–940, New York, NY, USA, 2008. ACM.

Contents

[RSTE] Problem Definition

- Two central elements
  - **User-Item Rating Matrix** (Fig. 1(b))
  - **Social Trust Graph** (Fig. 1(a))
    - Normally, the trust relations in the online trust network are **explicitly stated** by online users.



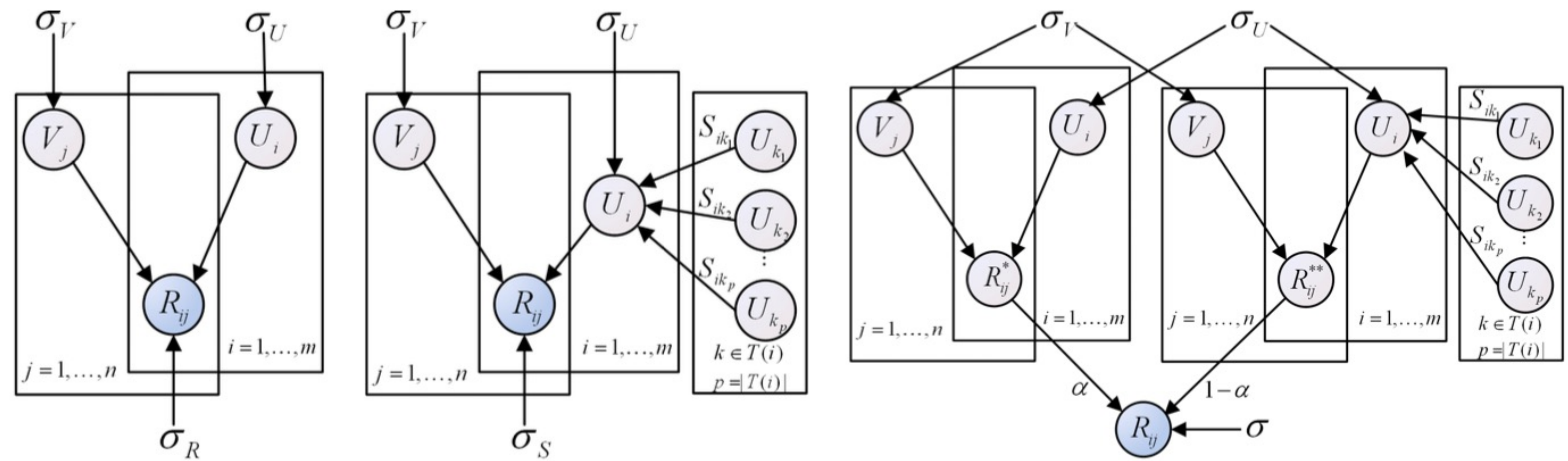
(a) Social Trust Graph

	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

(b) User-Item Rating Matrix

Figure 1: Example for Trust based Recommendation

[RSTE] Graphical Models



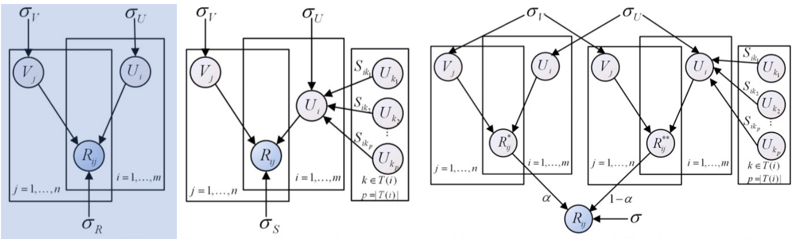
(a) Factorization of User-Item Matrix

(b) Recommendations by Trusted Friends

(c) Recommendations with Social Trust Ensemble

Figure 2: Graphical Models

[RSTE] User Features Learning



(a) Factorization of User-Item Matrix (b) Recommendations by Trusted Friends (c) Recommendations with Social Trust Ensemble

Figure 2: Graphical Models

- Conditional distribution

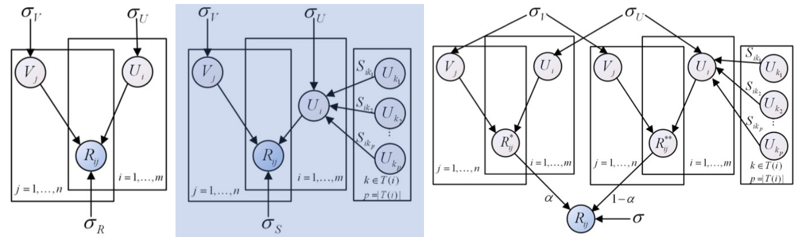
$$p(R|U,V,\sigma_R^2) = \prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathcal{N}\left(R_{ij}|g(U_i^T V_j),\sigma_R^2\right)\right]^{I_{ij}^R}, \quad (1)$$

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

Hence, through a Bayesian inference, we have

$$\begin{aligned} p(U,V|R,\sigma_R^2,\sigma_U^2,\sigma_V^2) &\propto p(R|U,V,\sigma_R^2)p(U|\sigma_U^2)p(V|\sigma_V^2) \\ &= \prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathcal{N}\left(R_{ij}|g(U_i^T V_j),\sigma_R^2\right)\right]^{I_{ij}^R} \\ &\times \prod_{i=1}^{m}\mathcal{N}(U_i|0,\sigma_U^2\mathbf{I}) \times \prod_{j=1}^{n}\mathcal{N}(V_j|0,\sigma_V^2\mathbf{I}). \quad (3) \end{aligned}$$

[RSTE] Rec. by Trusted Friends



(a) Factorization of User-Item Matrix (b) Recommendations by Trusted Friends (c) Recommendations with Social Trust Ensemble

Figure 2: Graphical Models

- Recommendations purely based on the trusted friends' tastes.

$$\widehat{R}_{ik} = \frac{\sum\limits_{j \in \mathcal{T}(i)} R_{jk} S_{ij}}{|\mathcal{T}(i)|}, \tag{4}$$

where $\widehat{R}_{ik}$ is the prediction of the rating that user u_i would give item v_j, R_{jk} is the score that user u_j gave item v_k, $\mathcal{T}(i)$ is the friends set that user u_i trusts and $|\mathcal{T}(i)|$ is the number of trusted friends of user u_i in the set $\mathcal{T}(i)$. $|\mathcal{T}(i)|$ can be merged into S_{ij} since it is the normalization term of trust scores. Hence, Eq. (4) can be simplified as

$$\widehat{R}_{ik} = \sum_{j \in \mathcal{T}(i)} R_{jk} S_{ij}. \tag{5}$$

[RSTE] Rec. by Trusted Friends

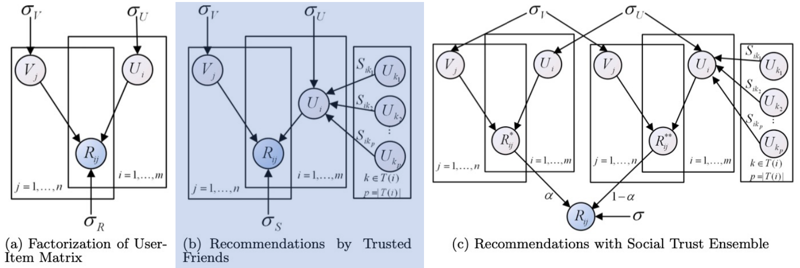


Figure 2: Graphical Models

- Example)

	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

(b) User-Item Rating Matrix

	1	2	3	4	5
1	0	0.6	0	1.0	0
2	0	0	0.9	0	1.0
3	0.2	0	0	0	0.8
4	0.8	0	0.5	0	0.6
5	0	0	0	0	0

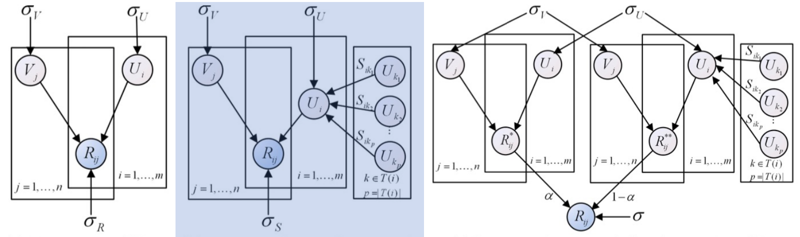
Social Trust Matrix

	1	2	3	4	5
1	0	0.3	0	0.5	0
2	0	0	0.45	0	0.5
3	0.1	0	0	0	0.4
4	0.4	0	0.25	0	0.3
5	0	0	0	0	0

Normalized Social Trust Matrix

$$R_{1:} = RS_{i:} = \begin{pmatrix} 0 & 4 & 0 & 5 & 0 \\ 5 & 0 & 0 & 0 & 5 \\ 2 & 0 & 2 & 0 & 5 \\ 0 & 3 & 0 & 3 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 0 & 4 & 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 0.3 \\ 0 \\ 0.5 \\ 0 \end{pmatrix} = \begin{pmatrix} 0+1.2+0+0.5+0 = 3.7 \\ \cancel{0+0+0+0+0=0} \\ \cancel{0+0+0+0+0=0} \\ 0+0.9+0+1.5+0 = 2.4 \\ \cancel{0+0+0+0+0=0} \\ 0+1.2+0+0+0 = 0.4 \end{pmatrix} = \begin{pmatrix} 3.7 \\ \cancel{0} \\ \cancel{0} \\ 2.4 \\ \cancel{0} \\ 0.4 \end{pmatrix}$$

[RSTE] Rec. by Trusted Friends



(a) Factorization of User-Item Matrix (b) Recommendations by Trusted Friends (c) Recommendations with Social Trust Ensemble

Figure 2: Graphical Models

- Conditional distribution

$$p(R|S,U,V,\sigma_R^2) = \prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathcal{N}\left(R_{ij}|g(\sum_{k\in\mathcal{T}(i)} S_{ik}U_k^T V_j),\sigma_S^2\right)\right]^{I_{ij}^R}, \quad (8)$$

$$p(U,V|R,S,\sigma_S^2,\sigma_U^2,\sigma_V^2) \propto p(R|S,U,V,\sigma_S^2)p(U|S,\sigma_U^2)p(V|S,\sigma_V^2). \quad (9)$$

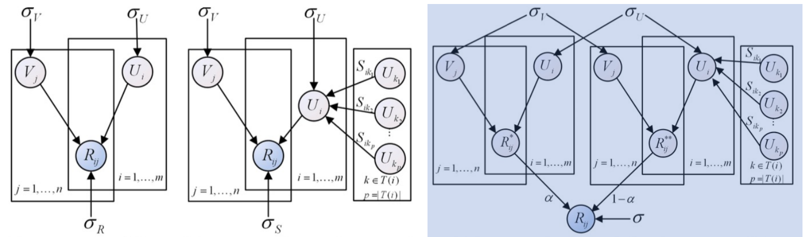
U,V,S는 서로 독립

$$p(U,V|R,S,\sigma_S^2,\sigma_U^2,\sigma_V^2) \propto p(R|S,U,V,\sigma_S^2)p(U|\sigma_U^2)p(V|\sigma_V^2),$$

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

$$\times \prod_{i=1}^{m}\mathcal{N}(U_i|0,\sigma_U^2\mathbf{I}) \times \prod_{j=1}^{n}\mathcal{N}(V_j|0,\sigma_V^2\mathbf{I}). \quad (10)$$

[RSTE] Social Trust **Ensemble**



(a) Factorization of User-Item Matrix (b) Recommendations by Trusted Friends (c) Recommendations with Social Trust Ensemble

Figure 2: Graphical Models

- Ensemble Formula

Target User i Trusted friends

$$
\begin{aligned}
&p(U, V | R, S, \sigma^2, \sigma_U^2, \sigma_V^2) \\
&= \prod_{i=1}^{m} \prod_{j=1}^{n} \left[\mathcal{N}\left(R_{ij} | g(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j), \sigma^2 \right) \right]^{I_{ij}^R} \\
&\times \prod_{i=1}^{m} \mathcal{N}(U_i | 0, \sigma_U^2 \mathbf{I}) \times \prod_{j=1}^{n} \mathcal{N}(V_j | 0, \sigma_V^2 \mathbf{I}). \qquad (11)
\end{aligned}
$$

[RSTE] Social Trust **Ensemble**

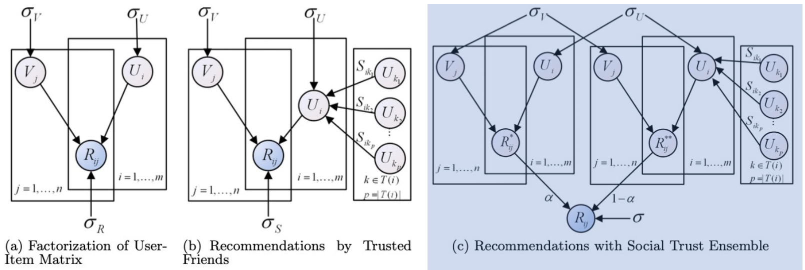


Figure 2: Graphical Models

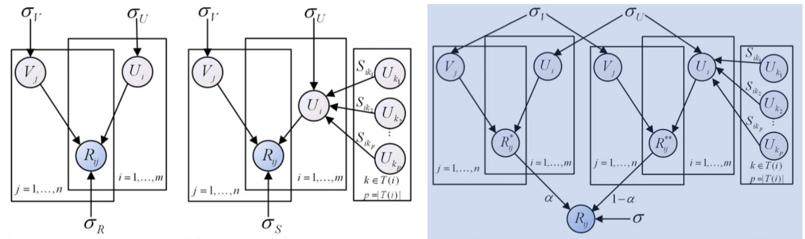
- Objective Function

$$\ln p(U,V|R,S,\sigma^2,\sigma_U^2,\sigma_V^2) =$$
$$-\frac{1}{2\sigma^2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}^R (R_{ij} - g(\alpha U_i^T V_j + (1-\alpha)\sum_{k\in\mathcal{T}(i)} S_{ik} U_k^T V_j))^2$$
$$-\frac{1}{2\sigma_U^2}\sum_{i=1}^{m} U_i^T U_i - \frac{1}{2\sigma_V^2}\sum_{j=1}^{n} V_j^T V_j$$
$$-\frac{1}{2}(\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}^R)\ln\sigma^2 - \frac{1}{2}(ml\ln\sigma_U^2 + nl\ln\sigma_V^2) + \mathcal{C}, \qquad (12)$$

$$\mathcal{L}(R,S,U,V)$$
$$= \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}^R (R_{ij} - g(\alpha U_i^T V_j + (1-\alpha)\sum_{k\in\mathcal{T}(i)} S_{ik} U_k^T V_j))^2$$
$$+ \frac{\lambda_U}{2}\|U\|_F^2 + \frac{\lambda_V}{2}\|V\|_F^2, \qquad (13)$$

where $\lambda_U = \sigma^2/\sigma_U^2$, $\lambda_V = \sigma^2/\sigma_V^2$, and $\|\cdot\|_F^2$ denotes the Frobenius norm.

[RSTE] Social Trust **Ensemble**



(a) Factorization of User-Item Matrix (b) Recommendations by Trusted Friends (c) Recommendations with Social Trust Ensemble

Figure 2: Graphical Models

- Gradient Descent in U_i, V_j.

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

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

- Predicted rate of User i Item j.

$$\widehat{R_{ij}} = U_i^T V_j$$

Contents

- Challenges in recommender system
- Social trust graph?
- Related works
- RSTE
- **Empirical analysis**
- Conclusions, future work and limitation

Empirical Analysis

- Epinions Dataset
  - Number of Users: 51,670
  - Number of Items: 83,509
  - Total number of ratings : 631,064
  - Density of the user-item rating matrix is less than 0.015%

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

Statistics	User	Item
Max. Num. of Ratings	1960	7082
Avg. Num. of Ratings	12.21	7.56

Table 2: Statistics of Social Trust Network of Epinions

Statistics	Trust per User	Be Trusted per User
Max. Num.	1763	2443
Avg. Num.	9.91	9.91

Empirical Analysis

$$p(U, V|R, S, \sigma^2, \sigma_U^2, \sigma_V^2)$$
$$= \prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathcal{N}\left(R_{ij}|g(\alpha U_i^T V_j + (1-\alpha)\sum_{k\in\mathcal{T}(i)} S_{ik}U_k^T V_j), \sigma^2\right)\right]^{I_{ij}^R}$$
$$\times \prod_{i=1}^{m}\mathcal{N}(U_i|0, \sigma_U^2\mathbf{I}) \times \prod_{j=1}^{n}\mathcal{N}(V_j|0, \sigma_V^2\mathbf{I}). \quad (11)$$

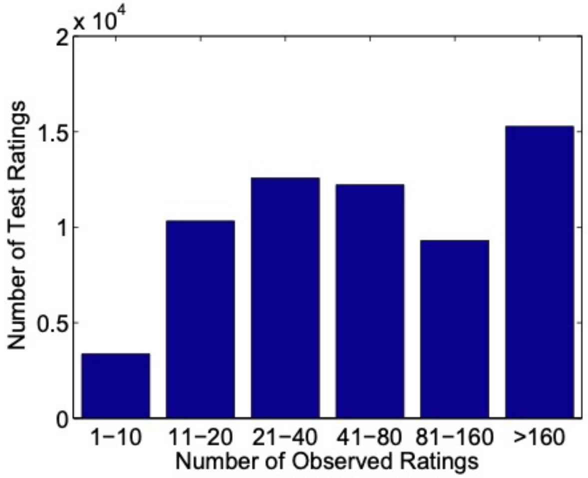
- PMF[19]
  - **Only uses user-item matrix**
  - Based on probabilistic matrix factorization
  - $\alpha = 1$
- Trust
  - **Purely uses trusted friend's tastes.**
  - $\alpha = 0$
- SoRec[13]
  - Factorizes the user-item rating matrix and users' social trust network by **sharing the same user latent space**.

Empirical Analysis

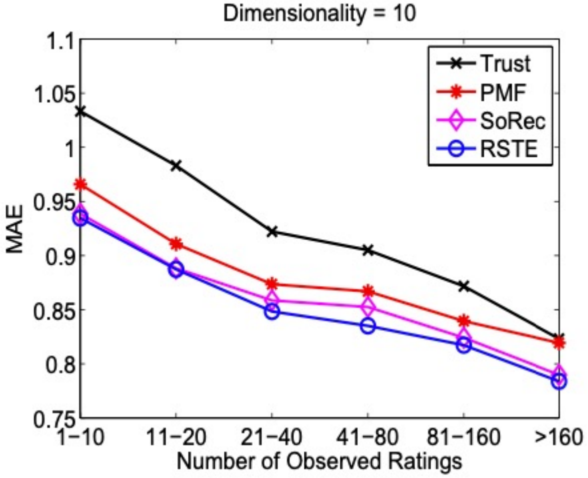
- Performance Comparison on Different Users.

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

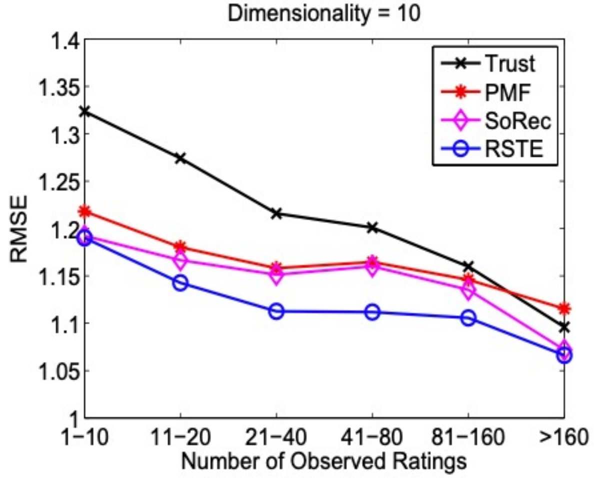
Training Data	Metrics	Dimensionality = 5				Dimensionality = 10			
		Trust	PMF	SoRec	RSTE	Trust	PMF	SoRec	RSTE
90%	MAE	0.9054	0.8676	0.8442	**0.8377**	0.9039	0.8651	0.8404	**0.8367**
	RMSE	1.1959	1.1575	1.1333	**1.1109**	1.1917	1.1544	1.1293	**1.1094**
80%	MAE	0.9221	0.8951	0.8638	**0.8594**	0.9215	0.8886	0.8580	**0.8537**
	RMSE	1.2140	1.1826	1.1530	**1.1346**	1.2132	1.1760	1.1492	**1.1256**



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



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



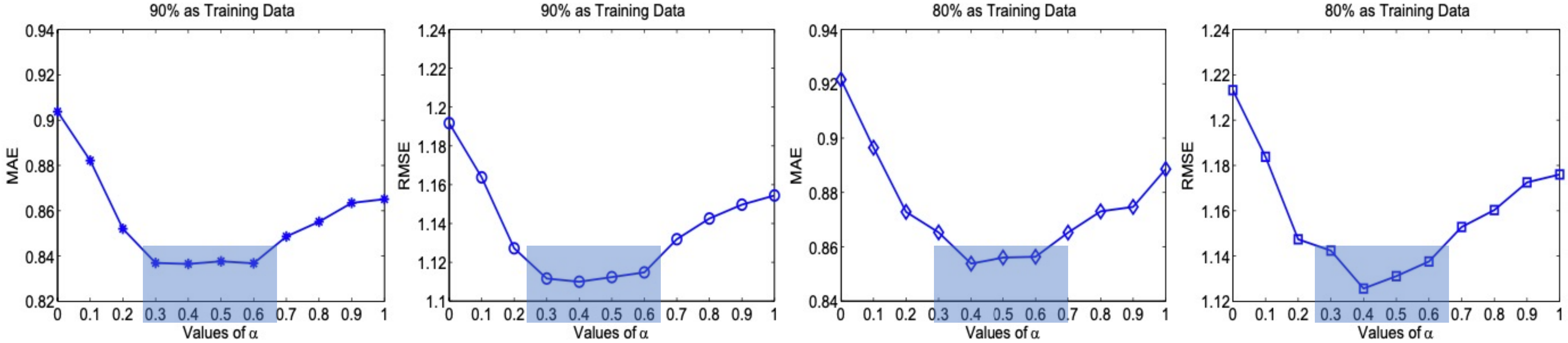
(c) RMSE Comparison on Different User Rating Scales (90% as Training Data)

Figure 3: Performance Comparison on Different Users

Empirical Analysis

$$p(U,V|R,S,\sigma^2,\sigma_U^2,\sigma_V^2)$$
$$= \prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathcal{N}\left(R_{ij}|g(\alpha U_i^T V_j + (1-\alpha)\sum_{k\in\mathcal{T}(i)} S_{ik}U_k^T V_j),\sigma^2\right)\right]^{I_{ij}^R}$$
$$\times \prod_{i=1}^{m}\mathcal{N}(U_i|0,\sigma_U^2\mathbf{I}) \times \prod_{j=1}^{n}\mathcal{N}(V_j|0,\sigma_V^2\mathbf{I}). \quad (11)$$

- Impact of Parameter α.



(a) (b) (c) (d)

Figure 4: Impact of Parameter α (Dimensionality = 10)

Empirical Analysis

$$p(U,V|R,S,\sigma^2,\sigma_U^2,\sigma_V^2)$$
$$= \prod_{i=1}^{m}\prod_{j=1}^{n}\left[\mathcal{N}\left(R_{ij}|g(\alpha U_i^T V_j + (1-\alpha)\sum_{k\in\mathcal{T}(i)} S_{ik}U_k^T V_j),\sigma^2\right)\right]^{I_{ij}^R}$$
$$\times \prod_{i=1}^{m}\mathcal{N}(U_i|0,\sigma_U^2\mathbf{I}) \times \prod_{j=1}^{n}\mathcal{N}(V_j|0,\sigma_V^2\mathbf{I}). \quad (11)$$

- Training Efficiency

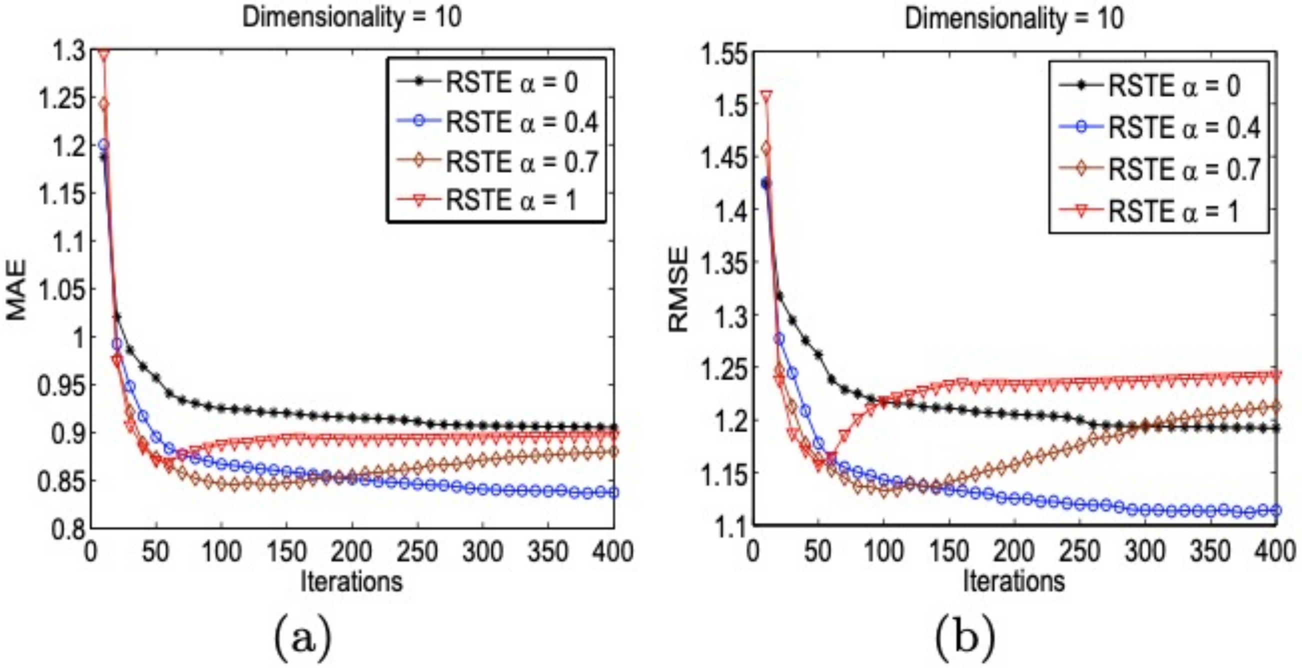


Figure 5: Efficiency Analysis (90% as Training Data)

Contents

- Challenges in recommender system
- Social trust graph?
- Related works
- RSTE
- Empirical analysis
- **Conclusions, future work and limitation**

Conclusions and Future work and Limitation

- Novel, effective and efficient probabilistic matrix factorization frame work for the recommender system employing social trust graph.

- Future work
  - Utilize User's dis-trust graph

- Limitations
  - It is difficult to get an explicit social graph. (추천인 / 친구 입력?)
  - Social graph could be sparse. ('cold-start problem')
  - My girlfriend won't buy gym equipment for our house…☹ ('misleading problem')