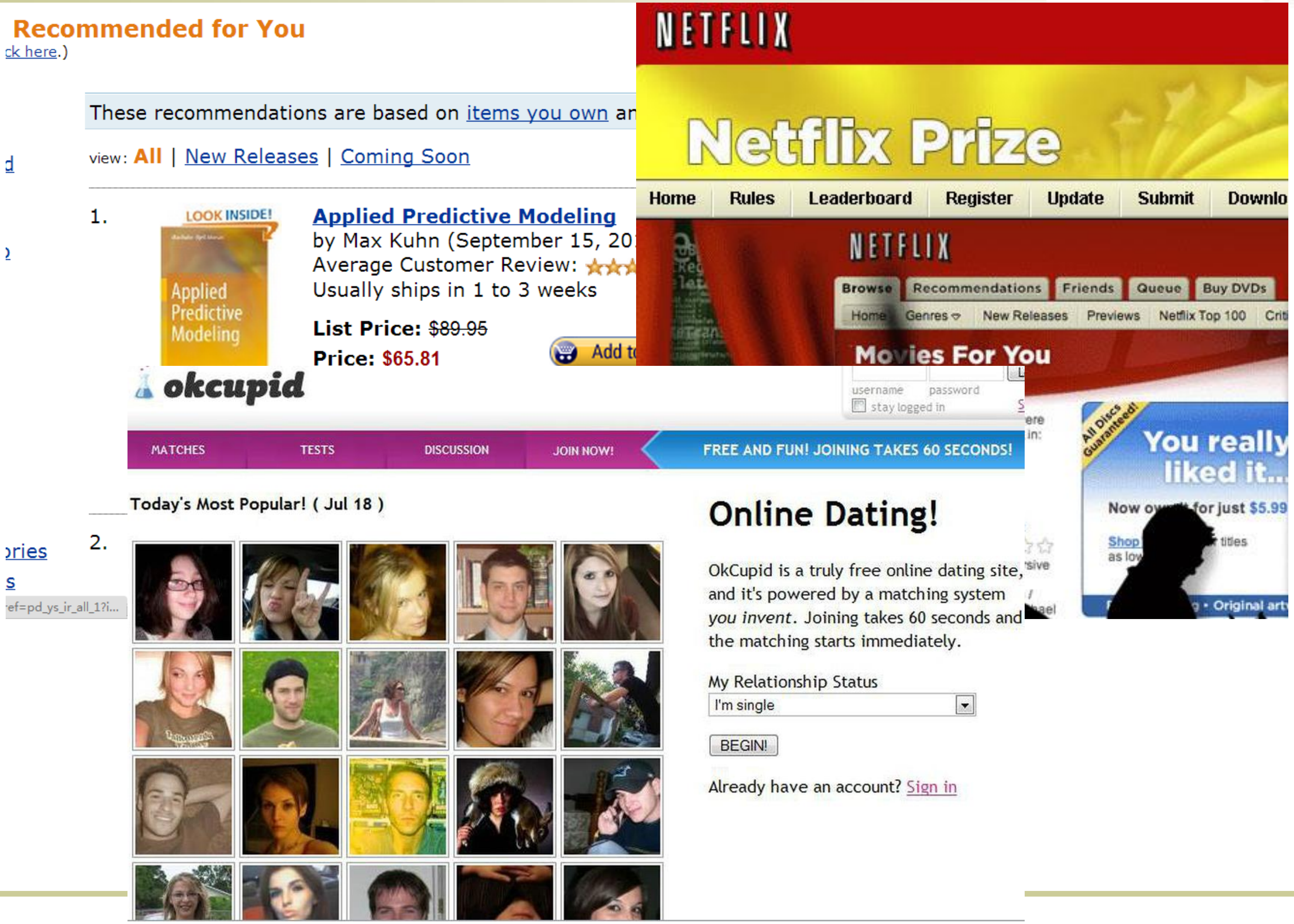
南京大學
NANJING UNIVERSITY

Combining Ratings, Social Relations, and Reviews for Recommendation

<u>Guang-Neng Hu</u>, Xin-Yu Dai,
Yunya Song (HKBU),
Shu-Jian Huang, Jia-Jun Chen

RSs are Ubiquitous

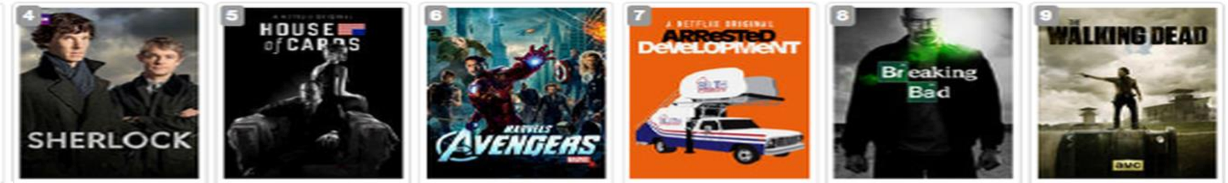
- Books at Amazon
- Movies at Netflix
- People at OkCupid



Recommendation as Rating Prediction

- *Rec*: Users x Items
  → Ratings

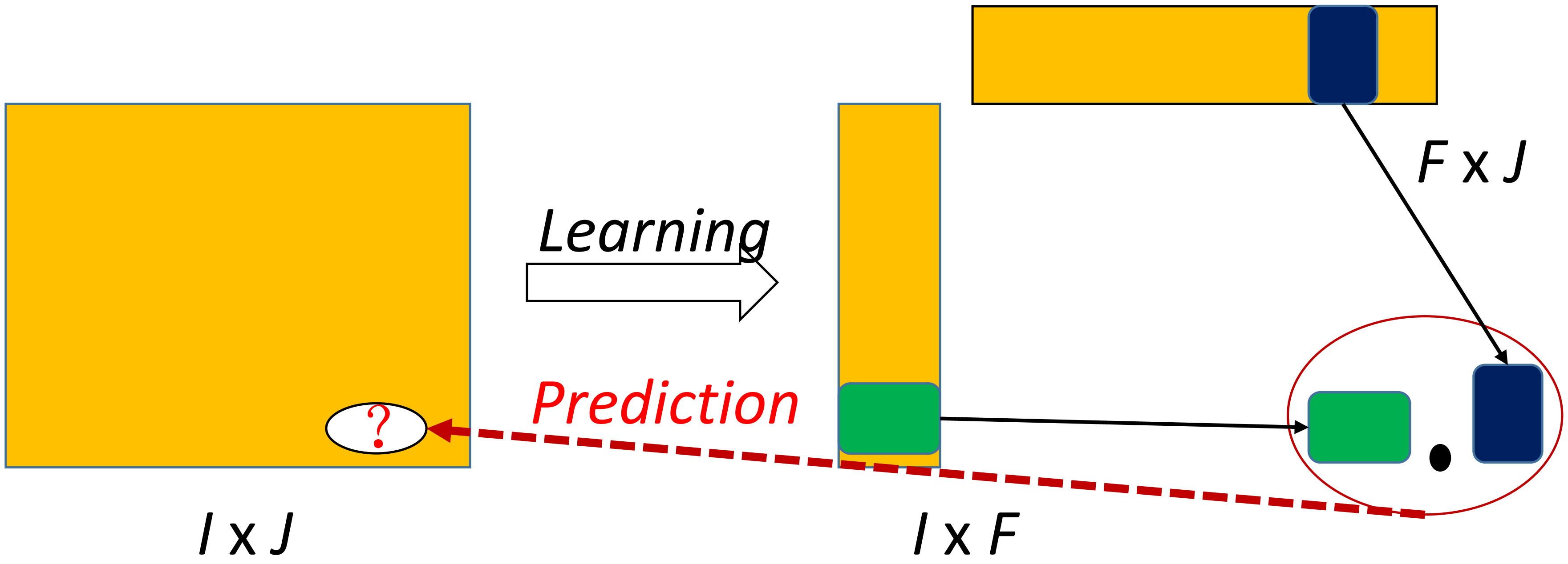
- Predict unknown ratings from observed data



2	?	2	?	5	?
5	?	4	?	?	1
?	?	5	?	2	?
?	1	?	5	?	?
?	5	?	1	?	4

Typical Model: Probabilistic Matrix Factorization (PMF)

- Low dimensional representations of users and of items



I: #Users, *J*: #Items

F: dimensionality of latent factors

Salakhutdinov & Mnih, Probabilistic matrix factorization, NIPS 2008

Issues of PMF

- Sparse rating matrix, e.g.,
  - Epinions: 0.022%
  - Ciao: 0.11%
- Cold-start users & items
  - Have no or few ratings

Statistics	Epinions	Ciao
# of Users	49,454	7,340
# of Items	74,154	22,472
# of Ratings/Reviews	790,940	183,974
# of Social Relations	434,680	112,942
# of Words	2,246,837	28,874,000
Rating Density	0.00022	0.0011
Social Density	0.00018	0.0021
Ave. Words Per Item	30.3	1284.9

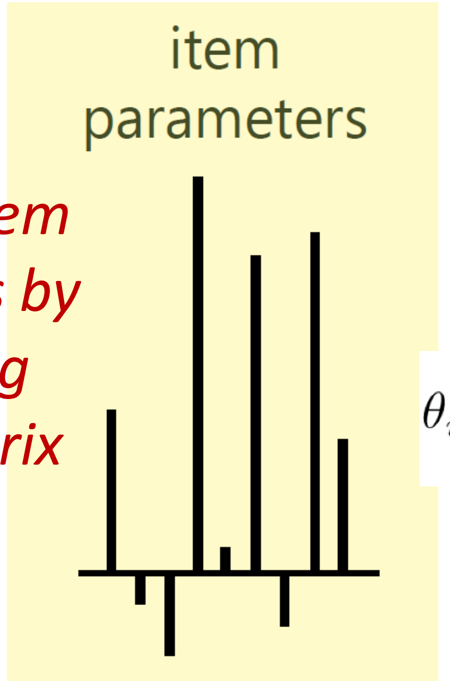
One Research Line to Address the Issues

- Topic MF: Integrating item reviews into ratings
  - Item reviews justify the ratings

One Research Line (cont')

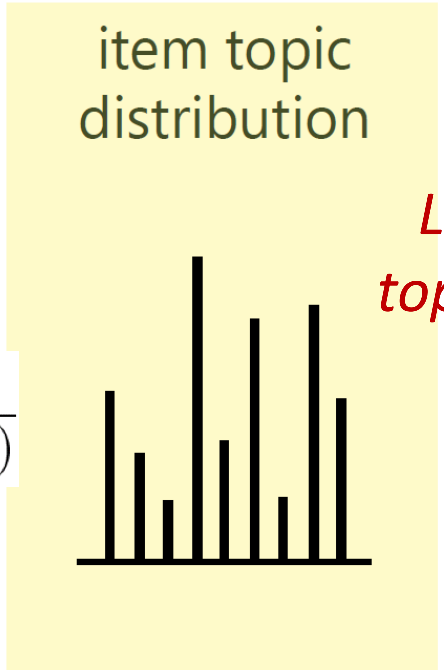
- Typical model: Hidden factors and hidden topics (HFT)

Learning item parameters by factorizing rating matrix



transform

$$\theta_{i,k} = \frac{\exp(\kappa \gamma_{i,k})}{\sum_{k'} \exp(\kappa \gamma_{i,k})}$$



Learning item topic distribution by topic modeling

$$\gamma_i \in \mathbb{R}^K$$

$$\theta_i \in \Delta^K \ (\text{i.e.,} \ \sum_k \theta_{i,k} = 1)$$

McAuley & Leskovec, Hidden factors and hidden topics, RecSys 2013

Another Research Line to Address the Issues

- Social MF: Integrating social relations into ratings
  - The rating behavior of users is influenced by their friends

Another Research Line (cont')

- Typical model: Local and global recommender (LOCABAL)

Exploiting global social context by user reputation

$$\min_{\mathbf{U},\mathbf{V},\mathbf{H}} \sum_{\langle u_i, v_j \rangle \in \mathcal{O}} w_i (\mathbf{R}_{ij} - \mathbf{u}_i^\top \mathbf{v}_j)^2$$

Exploiting ratings by learning latent representations of users and of items

$$+ \alpha \sum_{i=1}^{n} \sum_{u_k \in \mathcal{N}_i} (\mathbf{S}_{ik} - \mathbf{u}_i^\top \mathbf{H} \mathbf{u}_k)^2$$

Exploiting local social context by learning latent social representations

$$+ \lambda (\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2 + \|\mathbf{H}\|_F^2),$$

Tang et al., Exploiting local and global social context for recommendation, IJCAI 2013

Issues of Topic MF and Social MF

- Item reviews and social relations are both useful
  - Demonstrated by HFT and LOCABAL respectively

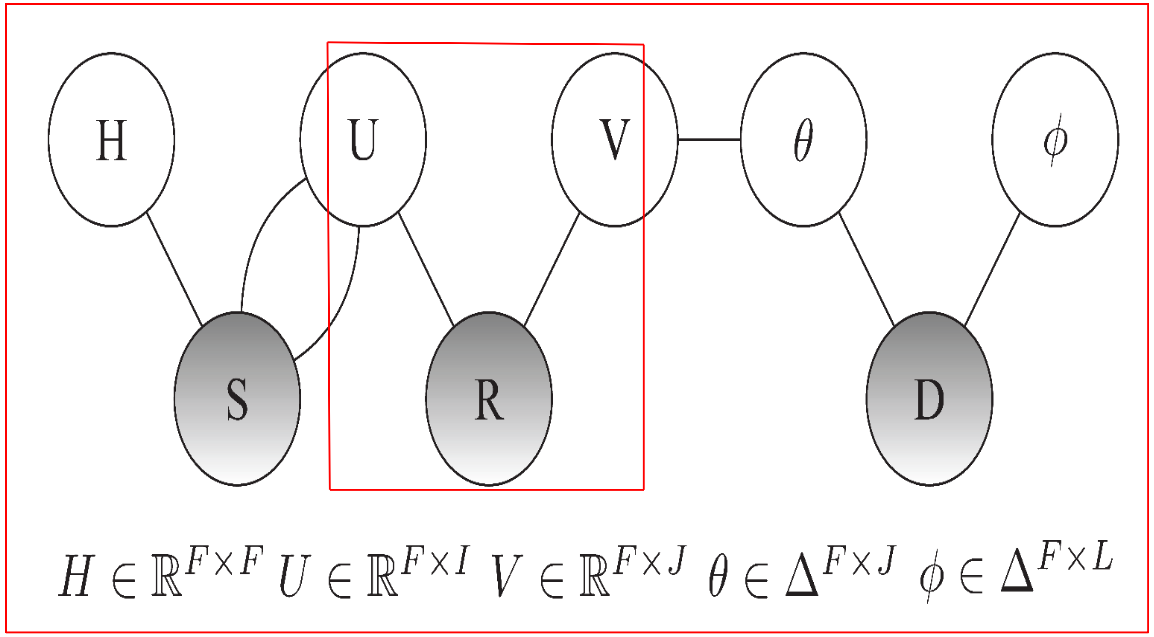
- Topic MF, e.g., HFT
  - ignores the social relations
- Social MF, e.g., LOCABAL
  - ignores the item reviews

Combining Ratings, Social Relations, and Reviews for Recommendation

- Item reviews and social relations are both useful for improving rating prediction

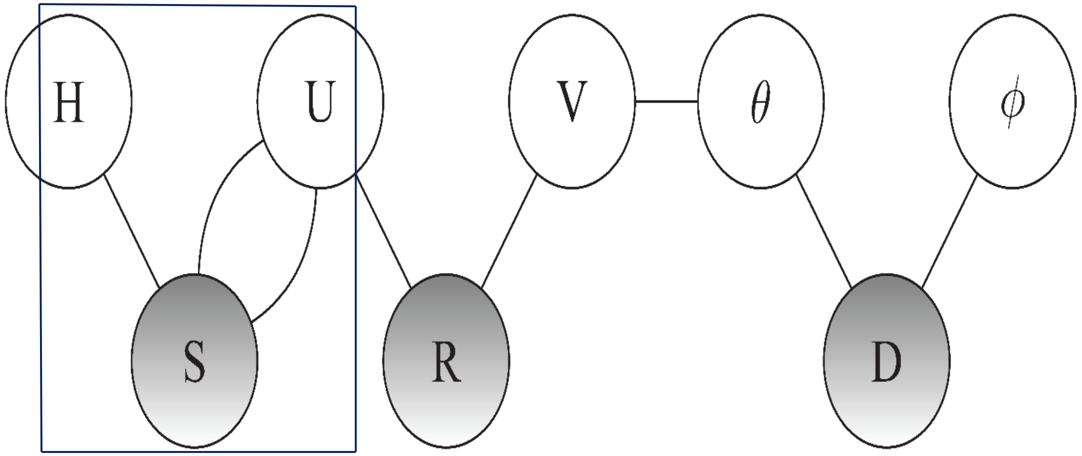
Modelling Three Kinds of Data Sources

- *Key*: connecting relations and reviews through ratings
  - For rating source, learning latent representations of users and of items



Modelling Three Kinds of Data Sources

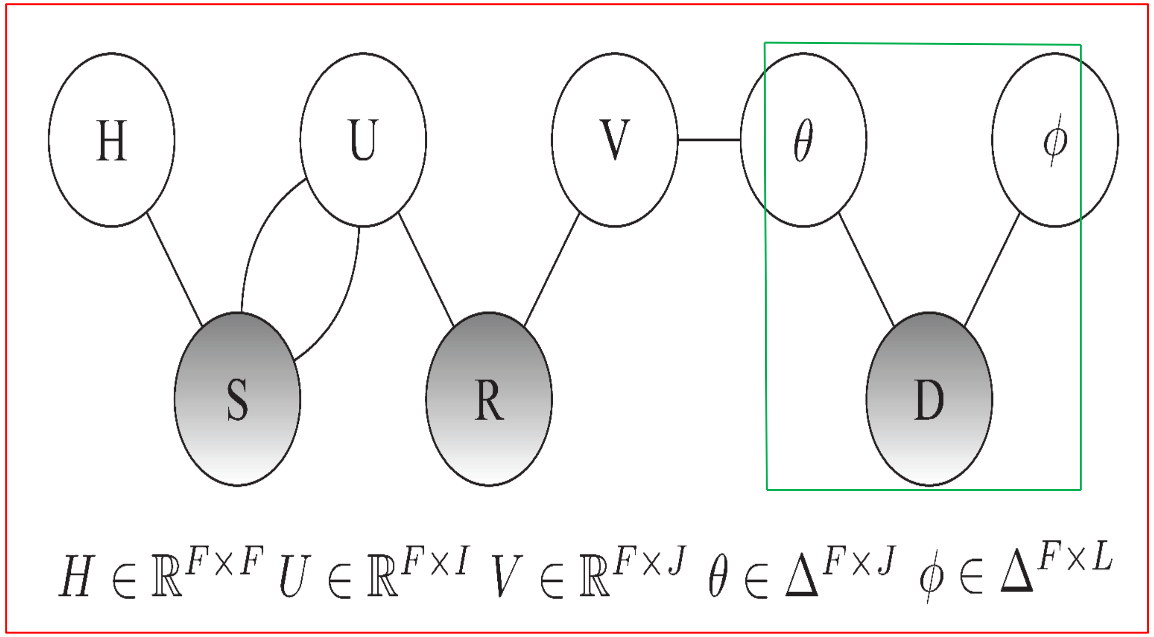
- *Key*: connecting relations and reviews through ratings
  - For rating source, learning latent representations of users and of items
  - For social relation source, learning latent social representations of users and their social relation matrix



$H \in \mathbb{R}^{F\times F}$ $U \in \mathbb{R}^{F\times I}$ $V \in \mathbb{R}^{F\times J}$ $\theta \in \Delta^{F\times J}$ $\phi \in \Delta^{F\times L}$

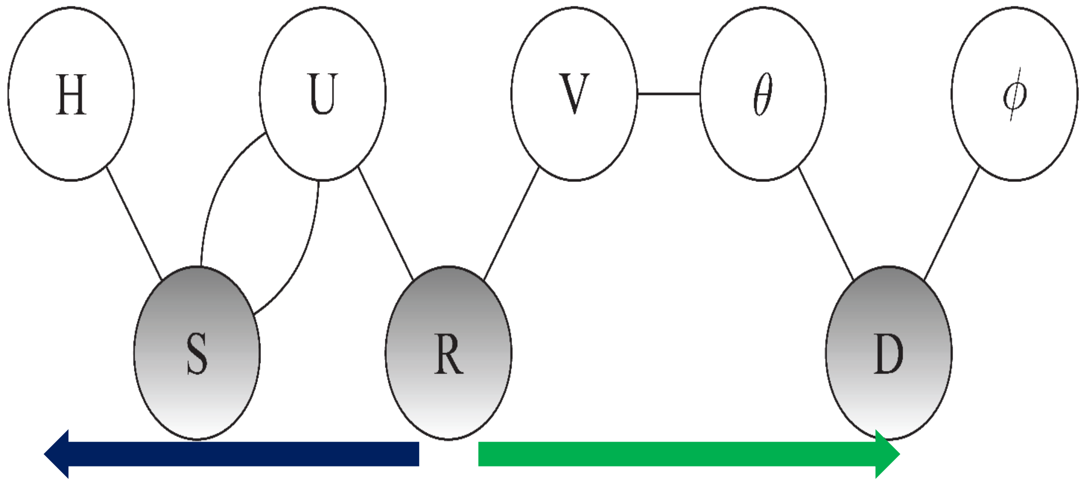
Modelling Three Kinds of Data Sources

- *Key*: connecting relations and reviews through ratings
  - For rating source, learning latent representations of users and of items
  - For social relation source, learning latent social representations of users and their social relation matrix
  - For item reviews, learning topic distributions (and word distributions)



Modelling Three Kinds of Data Sources

- *Key*: connecting relations and reviews through ratings
  - For rating source, learning latent representations of users and of items
  - For social relation source, learning latent social representations of users and their social relation matrix
  - For item reviews, learning topic distributions and word distributions



$H \in \mathbb{R}^{F \times F}$ $U \in \mathbb{R}^{F \times I}$ $V \in \mathbb{R}^{F \times J}$ $\theta \in \Delta^{F \times J}$ $\phi \in \Delta^{F \times L}$

MR3: A Model of Ratings, Reviews and Relations

$$\mathcal{L}(\Theta, \Phi, z, \kappa) \triangleq \sum_{R_{i,j} \neq 0} W_{i,j} \underbrace{\left(R_{i,j} - \hat{R}_{i,j}\right)^2}_{}$$

Exploiting ratings

$$- \lambda_{\text{rev}} \sum_{d=1}^{J} \sum_{n \in N_d} \left(\log \theta_{z_{d,n}} + \log \phi_{z_{d,n}, w_{d,n}}\right)$$

Exploiting reviews

$$+ \lambda_{\text{rel}} \sum_{T_{i,k} \neq 0} \boxed{C_{i,k}} \left(S_{i,k} - U_i^{\mathrm{T}} H U_k\right)^2 + \lambda \Omega(\Theta),$$

Exploiting social relations

here parameters $\Theta = \{U, V, H\}$ are associated with rat

nd social relations, parameters $\Phi = \{\theta, \phi\}$ associated

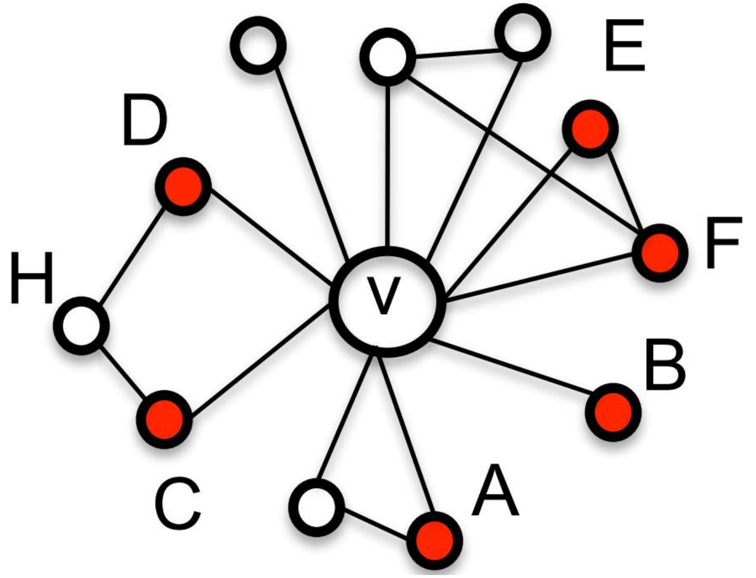
By-product of MR3: eSMF (extended Social MF)

- Graph structure of neighbors captures social influence locality, i.e., user behaviors are mainly influenced by close/direct friends in their ego networks

$$\min_{U,V,H} \sum_{R_{i,j} \neq 0} W_{i,j}(R_{i,j} - \hat{R}_{i,j})^2 + \lambda \sum_{T_{i,k} \neq 0} \boxed{C_{i,k}}(S_{i,k} - U_i^{\mathrm{T}} H U_k)^2 + \lambda\Omega(\Theta).$$

The trust values

$$C_{ik} = \sqrt{d_{u_k}^- / (d_{u_i}^+ + d_{u_k}^-)},$$



Ma et al., SoRec: Social Recommendation Using Probabilistic Matrix Factorization, CIKM 2008
Zhang et al., Social influence locality for modeling retweeting behaviors, IJCAI 2013

Learning Parameters

- Alternating two steps
  - Topic assignments $z_{d,n}$ for each word in reviews corpus are fixed; then we update the terms Θ, Φ, and κ by gradient descent
  - Parameters associated with reviews corpus θ and ϕ are fixed; then sample $z_{d,n}$ by iterating through all docs and each word within

$$\text{update } \Theta^{\text{new}}, \Phi^{\text{new}}, \kappa^{\text{new}} = \arg\min_{\Theta,\Phi,\kappa} \mathcal{L}(\Theta, \Phi, \kappa, z^{\text{old}}); \quad (1$$

$$\text{sample } \; z_{d,n}^{\text{new}} \text{ with probability } p(z_{d,n}^{\text{new}} = f) = \phi_{f,w_{d,n}}^{\text{new}}.$$

Gradient descent

- Alternating two steps
  - Topic assignments *zd,n* for each word in reviews corpus are fixed; then we update the terms Θ, Φ, and *κ* by gradient descent
  - Parameters associated with reviews corpus θ and ϕ are fixed; then sample *zd,n* by iterating through all docs and each word within

$$\frac{1}{2}\frac{\partial \mathcal{L}}{\partial U_i} = \sum_{j:R_{i,j}\neq 0} W_{i,j}(\hat{R}_{i,j} - R_{i,j})V_j + \lambda U_i + \lambda_{\text{rel}} \sum_{k:T_{k,i}\neq 0} C_{i,k}(U_k^{\mathrm{T}} H U_i - S_{i,k}) H^{\mathrm{T}} U_k + \lambda_{\text{rel}} \sum_{k:T_{i,k}\neq 0} C_{k,i}(U_i^{\mathrm{T}} H U_k - S_{i,k}) H U_k.$$

$$\frac{\partial \mathcal{L}}{\partial V_j} = 2\sum_{i:R_{i,j}\neq 0} W_{i,j}(\hat{R}_{i,j} - R_{i,j})U_i - \lambda_{\text{rev}}\kappa\left(M_j - \frac{m_j}{z_j}\exp\left(\kappa V_j\right)\right) + 2\lambda V_j.$$

$$\frac{1}{2}\frac{\partial \mathcal{L}}{\partial H} = \lambda_{\text{rel}} \sum_{T_{i,k}\neq 0} C_{i,k}(U_i^{\mathrm{T}} H U_k - S_{i,k}) U_i U_k^{\mathrm{T}} + \lambda H.$$

$$\frac{\partial \mathcal{L}}{\partial \psi_{fw}} = -\lambda_{\text{rev}}\left(M_{fw} - \frac{m_f}{z_f}\exp\left(\psi_{fw}\right)\right).$$

$$\frac{\partial \mathcal{L}}{\partial \kappa} = -\lambda_{\text{rev}} \sum_{j,f} V_{jf}\left(M_{jf} - \frac{m_j}{z_j}\exp\left(\kappa V_{jf}\right)\right).$$

Datasets

- Epinions and Ciao
  - http://www.public.asu.edu/~jtang20/

Statistics	Epinions	Ciao
# of Users	49,454	7,340
# of Items	74,154	22,472
# of Ratings/Reviews	790,940	183,974
# of Social Relations	434,680	112,942
# of Words	2,246,837	28,874,000
Rating Density	0.00022	0.0011
Social Density	0.00018	0.0021
Ave. Words Per Item	30.3	1284.9

Metric and Code

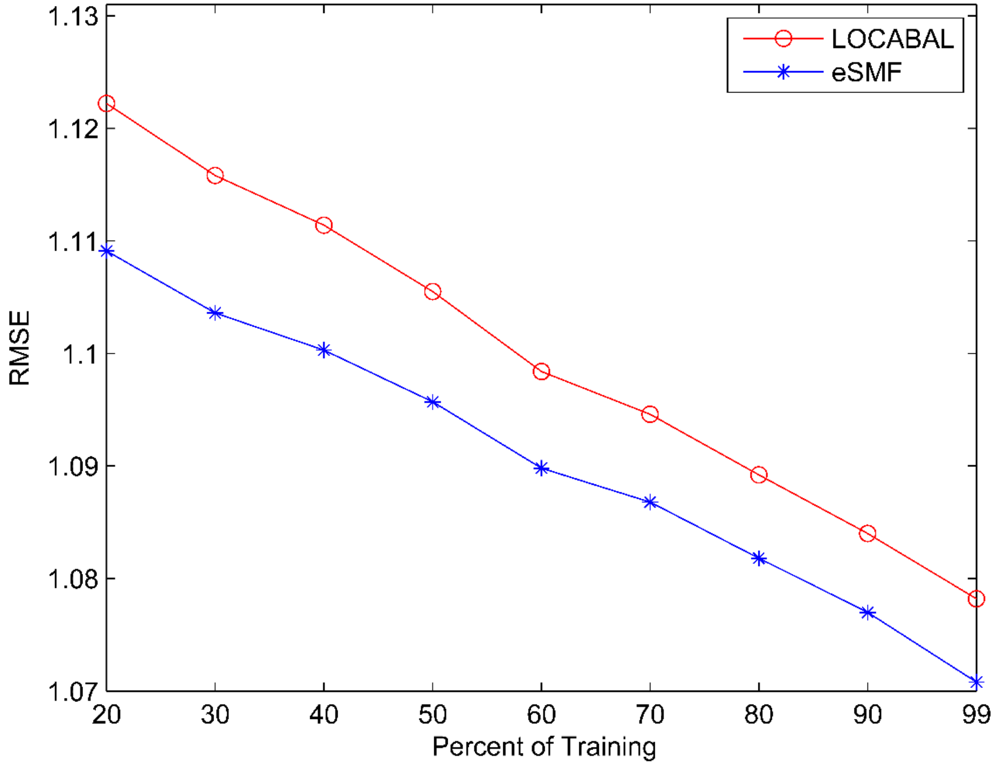
- RMSE (root-mean-square error)
  - The lower, the better

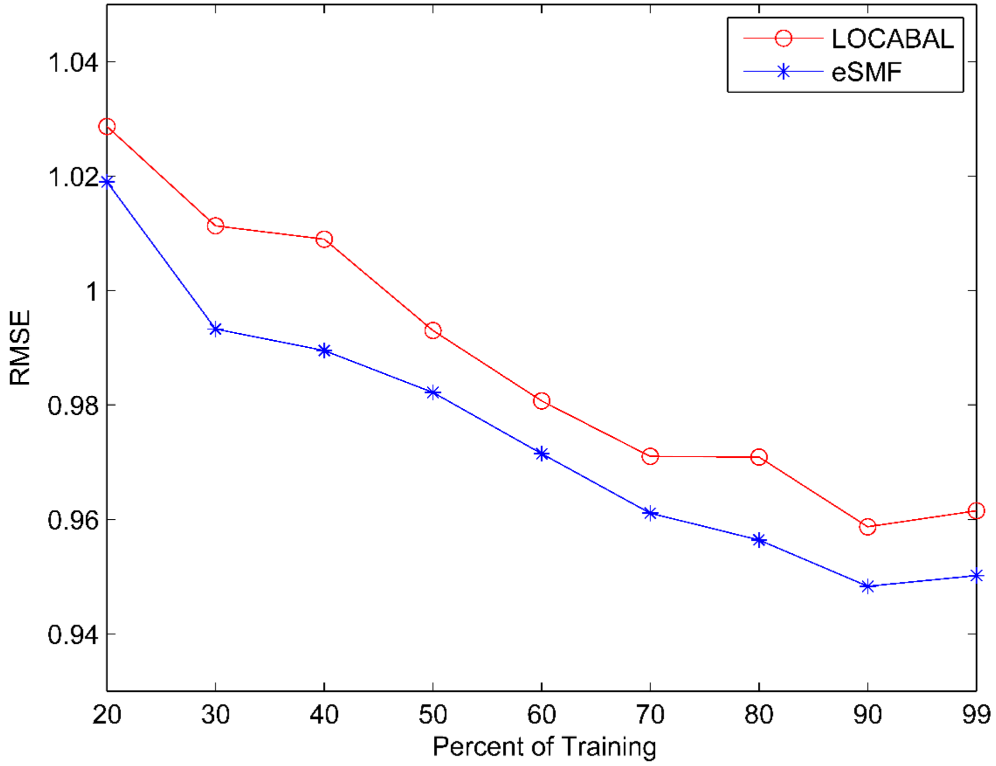
$$RMSE_{\mathcal{T}} = \sqrt{\sum_{(u_i,v_j)\in\mathcal{T}} (R_{i,j} - \hat{R}_{i,j})^2 \Big/ |\mathcal{T}|}$$

- PMF
  - http://www.cs.toronto.edu/~rsalakhu/BPMF.html
- HFT
  - http://cseweb.ucsd.edu/~jmcauley/

Comparing Social MF

- eSMF vs. LOCABAL





Left: Epinions; Right: Ciao

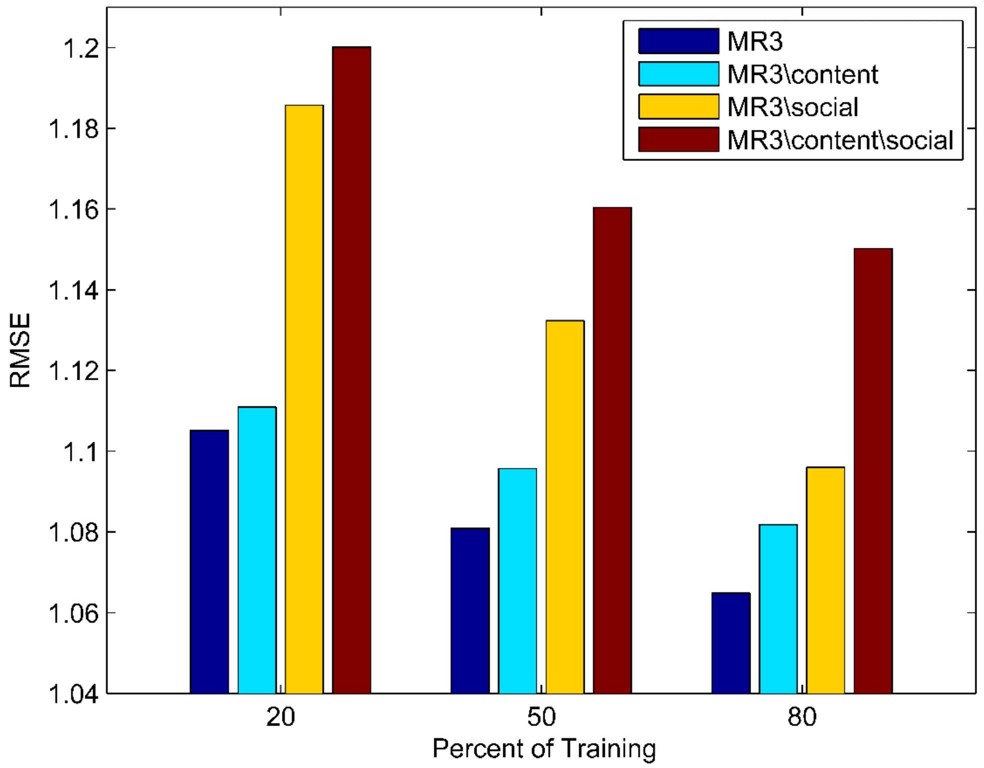
Comparing Different Recommender Systems

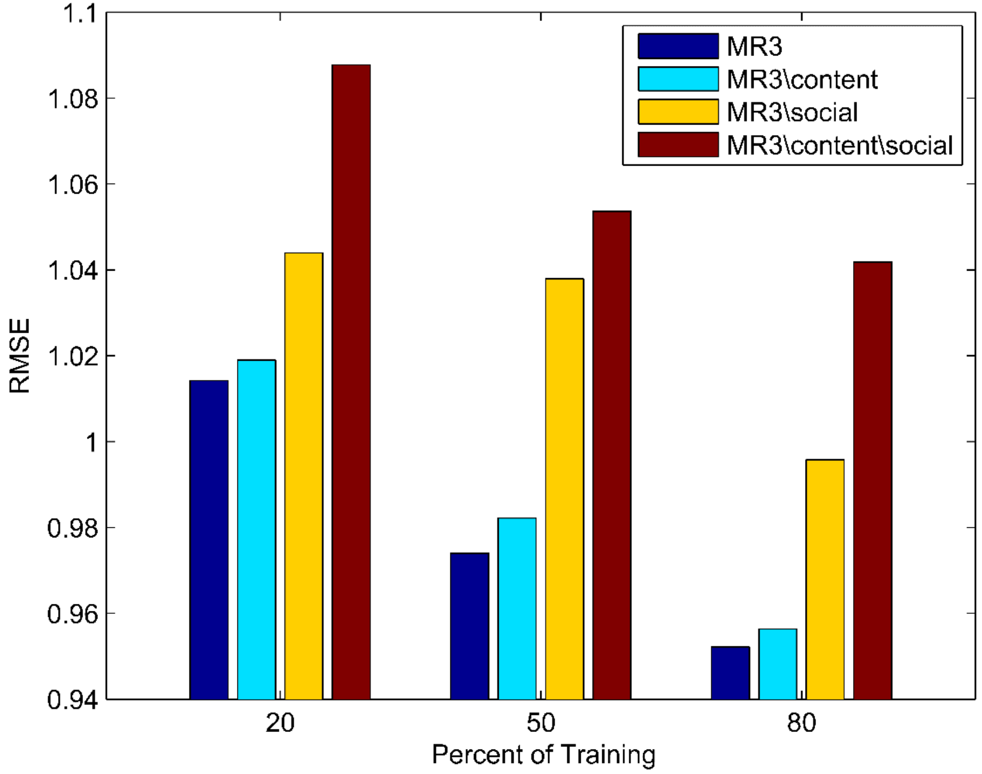
- MR3 vs. PMF, HFT, and LOCABAL (F = 10)

Datasets	Training	Methods					Improvement of MR3 vs.		
		Mean	PMF	HFT	LOCABAL	MR3	PMF	HFT	LOCABAL
Epinions	20%	1.2265	1.2001	1.1857	1.1222	1.1051	8.60%	7.29%	1.55%
	50%	1.2239	1.1604	1.1323	1.1055	1.0809	7.35%	4.76%	2.28%
	80%	1.2225	1.1502	1.0960	1.0892	1.0648	8.02%	2.93%	2.29%
	90%	1.2187	1.1484	1.0867	1.0840	1.0634	7.99%	2.19%	1.94%
Ciao	20%	1.1095	1.0877	1.0439	1.0287	1.0142	7.25%	2.93%	1.43%
	50%	1.0964	1.0536	1.0379	0.9930	0.9740	8.17%	6.56%	1.95%
	80%	1.0899	1.0418	0.9958	0.9709	0.9521	9.42%	4.59%	1.97%
	90%	1.0841	1.0391	0.9644	0.9587	0.9451	9.95%	2.04%	1.44%
Average							8.34%	4.16%	1.86%

Impact of Reviews and Social Relations

- MR3 compared with its three components (F = 10)

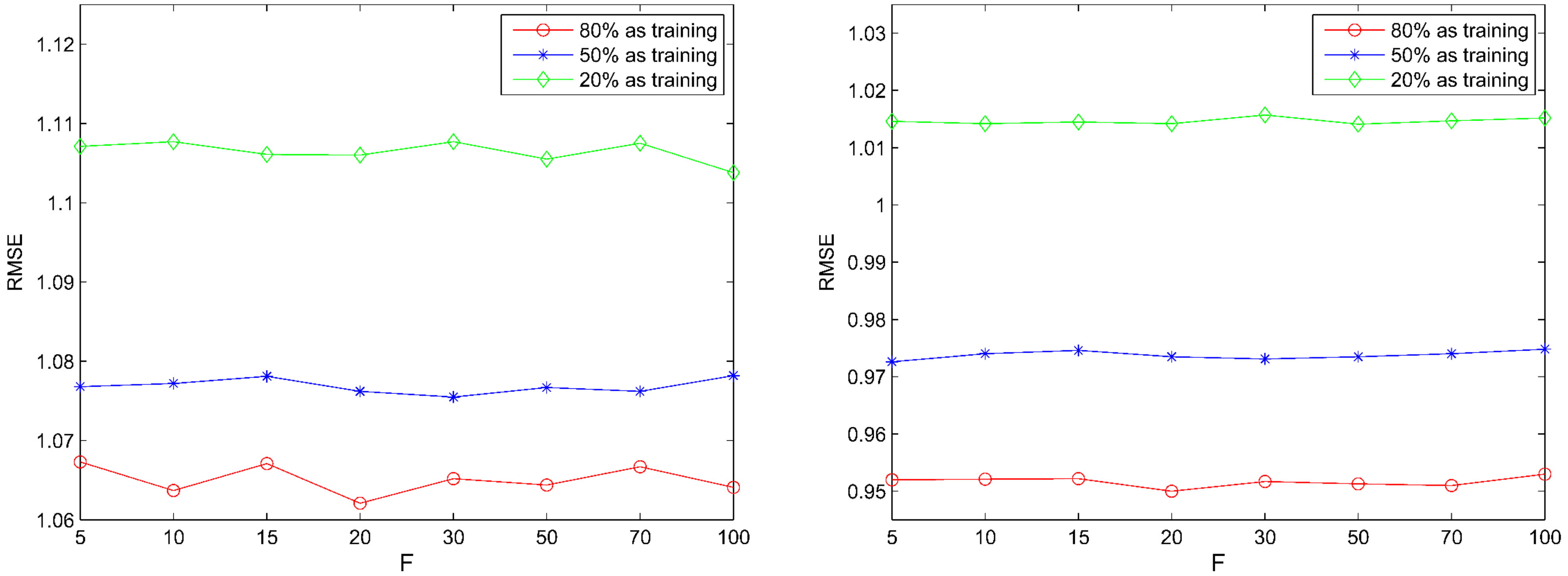




Left: Epinions; Right: Ciao

Sensitivity to Parameters

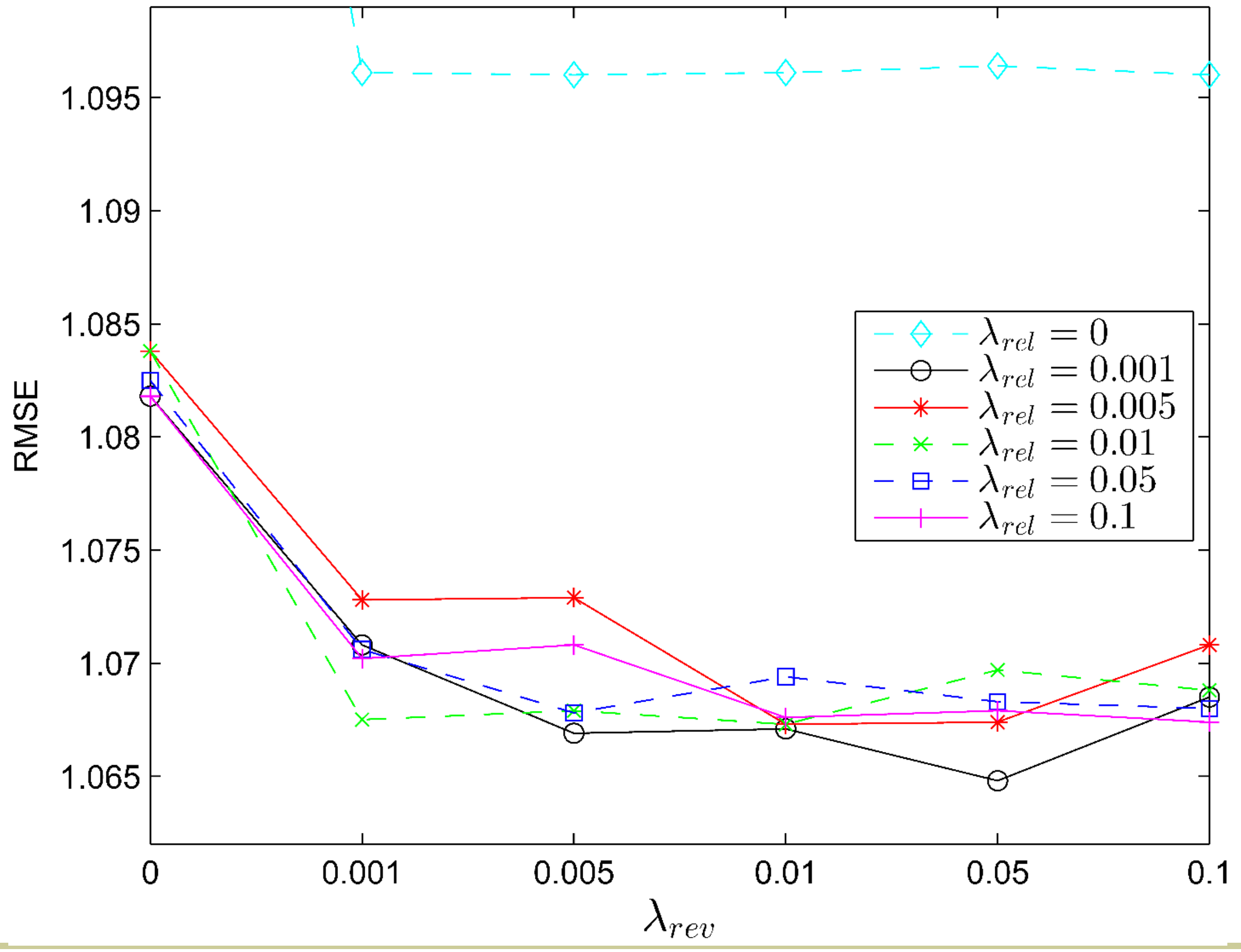
- F: the number of latent factors; Default: 10



Left: Epinions; Right: Ciao

Sensitivity to Parameters (cont')

- λ*rel*: controls the contribution from social *rel*ations
- λ*rev*: controls the contribution from *rev*iews
- Default: 0.001, 0.05



Conclusions

- A novel framework to exploit ratings, social relations, and reviews simultaneously for recommendation

- An advanced method to exploit ratings and social relations more tightly by capturing the graph structure of neighbors

- Significant improvements over the state of the art methods on the rating prediction task

Future Works

- Implicit feedback
- Temporal dynamics
- Number of hidden topics in reviews different from that of latent factors in ratings

References

- Salakhutdinov & Mnih, *Probabilistic matrix factorization*, NIPS 2008
- McAuley & Leskovec, *Hidden factors and hidden topics*, RecSys 2013
- Ma et al., *SoRec: Social Recommendation Using Probabilistic Matrix Factorization*, CIKM 2008
- Tang et al., *Exploiting local and global social context for recommendation*, IJCAI 2013
- Zhang et al., Social influence locality for modeling retweeting behaviors, IJCAI 2013
- **Hu et al., *A Synthetic Approach for Recommendation: Combining Ratings, Social Relations, and Reviews*, to appear in IJCAI 2015**

Thanks

Q&A