Data-Driven Behavioral Analytics:
Observations, Representations and Models

Meng Jiang (UIUC)
Peng Cui (Tsinghua)
Jiawei Han (UIUC)

http://www.meng-jiang.com/tutorial-cikm16.html

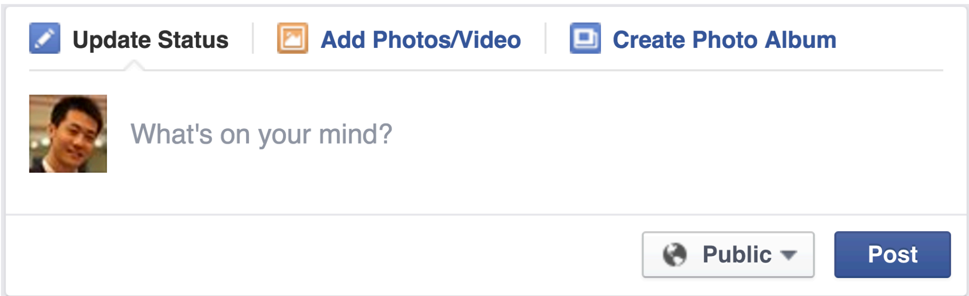
Tutorial in CIKM 2016, October 24, Indianapolis, IN

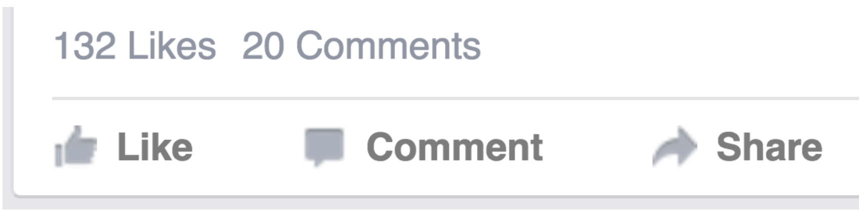
I. Mining behavior networks with social and spatiotemporal contexts

I.1. Behavior prediction and recommendation

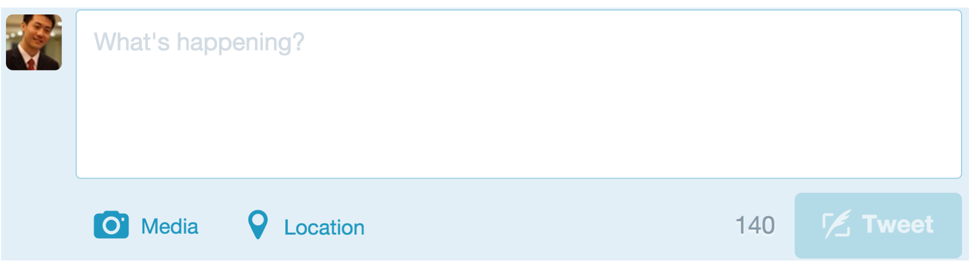
Behavior in Social Networks

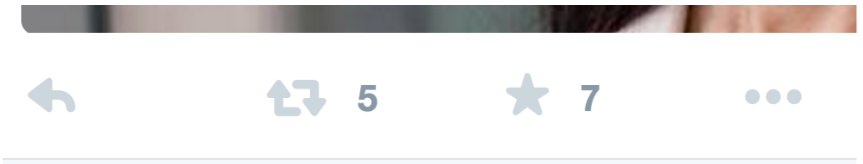
- **Facebook:** Post, Like, Comment, Share



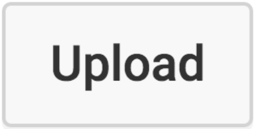


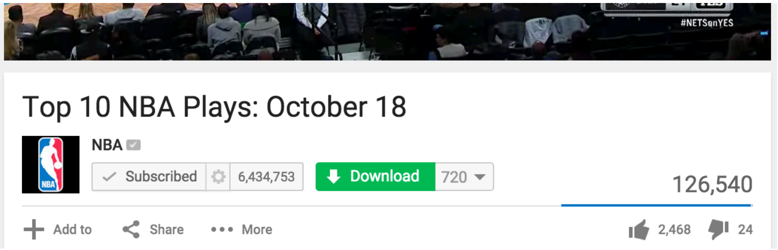
- **Twitter:** Post, Reply, Retweet, Favorite





- **YouTube:** Upload, Subscribe, Download, Share, Comment

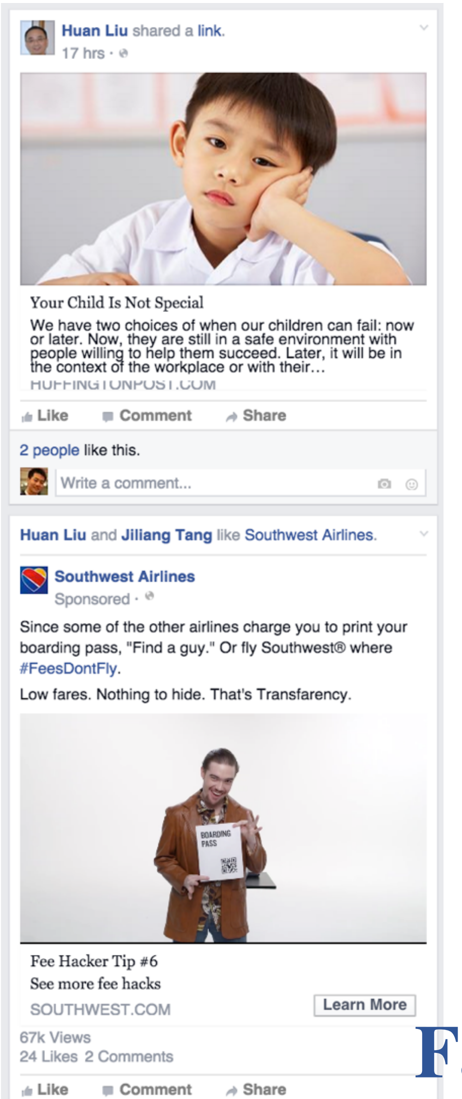


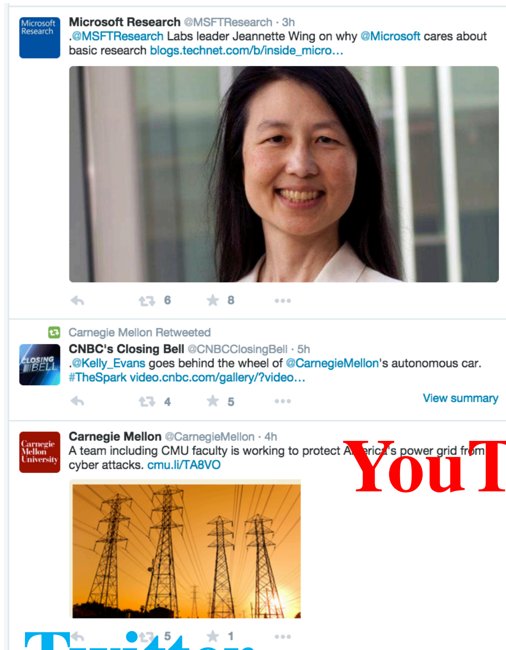


Behavior in Social Networks

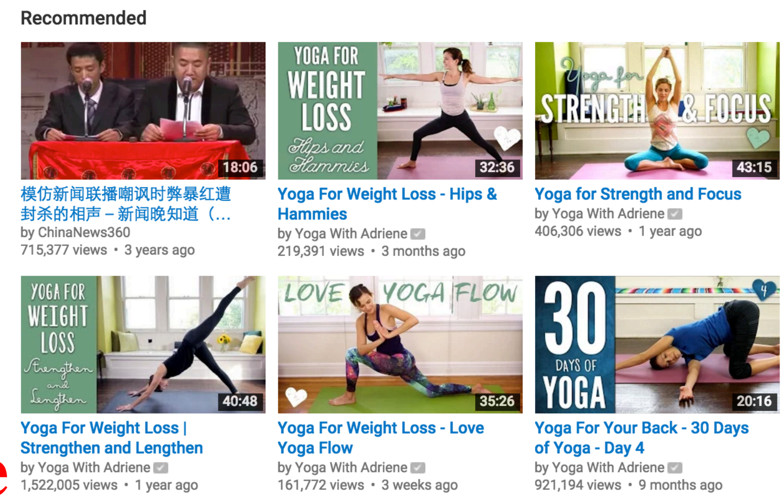
Like
Reply
Share
Favorite
Retweet
Comment
Subscribe
Download
Add to
Send
Pin it
Visit
……

Social Recommender Systems

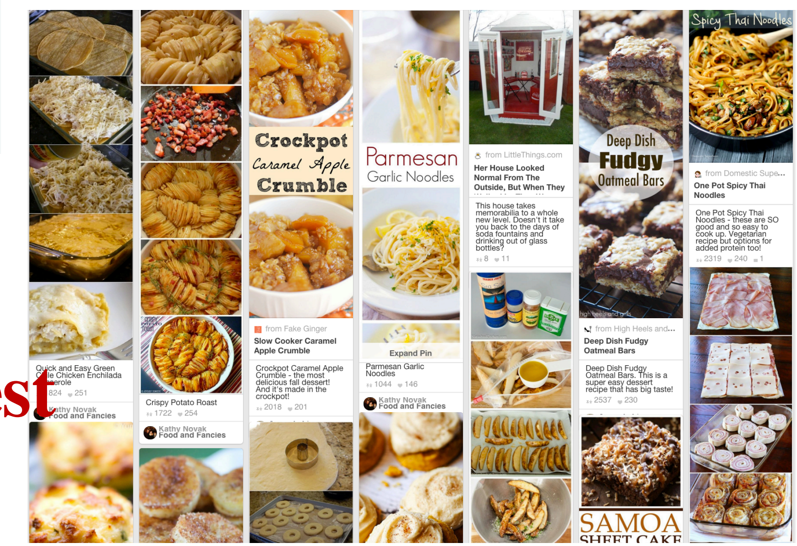




Twitter



YouTube



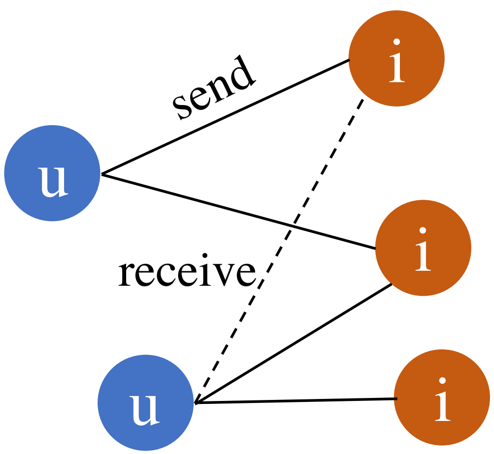
Pinterest

Facebook

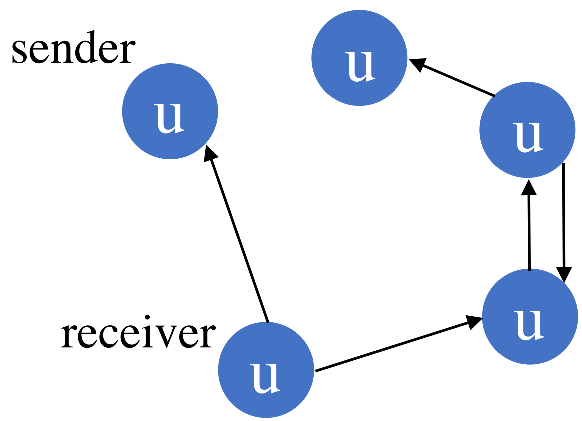
Social Recommender Systems

- April 20, 2011: **Tencent Weibo** visited Tsinghua University
  - Low *conversion rate* (< 6%): #retweets per feed request
  - Can we build a ***social recommender system***?
  - **Given**

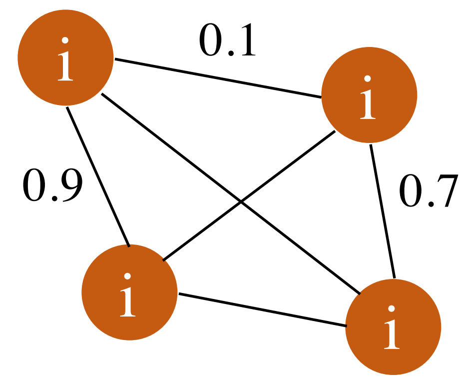
User-item behavior network



User-user social network



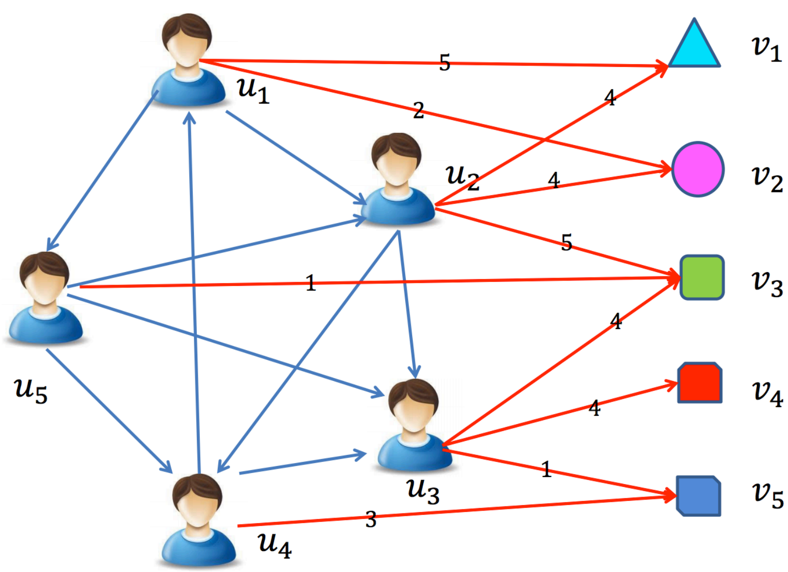
Content similarity (topic level) [Blei *et al.*]



- **Predict** which tweet/item a user will retweet.

Traditional Recommender Systems

- Assumed that users are independent and identically distributed (user-movie, user-book, *etc*.)



	v_1	v_2	v_3	v_4	v_5
u_1	5	?	2	?	?
u_2	4	4	5	?	?
u_3	?	?	4	4	1
u_4	?	?	?	?	3
u_5	?	?	1	?	?

Traditional Recommender Systems

- Content-based recommender (e.g., TF-IDF)
  - For textual information (e.g., news, documents)
  - *Limitation:* limited content analysis, over-specialization
- Collaborative filtering based recommender
  - Memory-based CF (e.g., PCC, similarity)
  - Model-based CF (e.g., factorization based)
  - *Limitation:* data sparsity, cold-start problem
- Hybrid recommender system

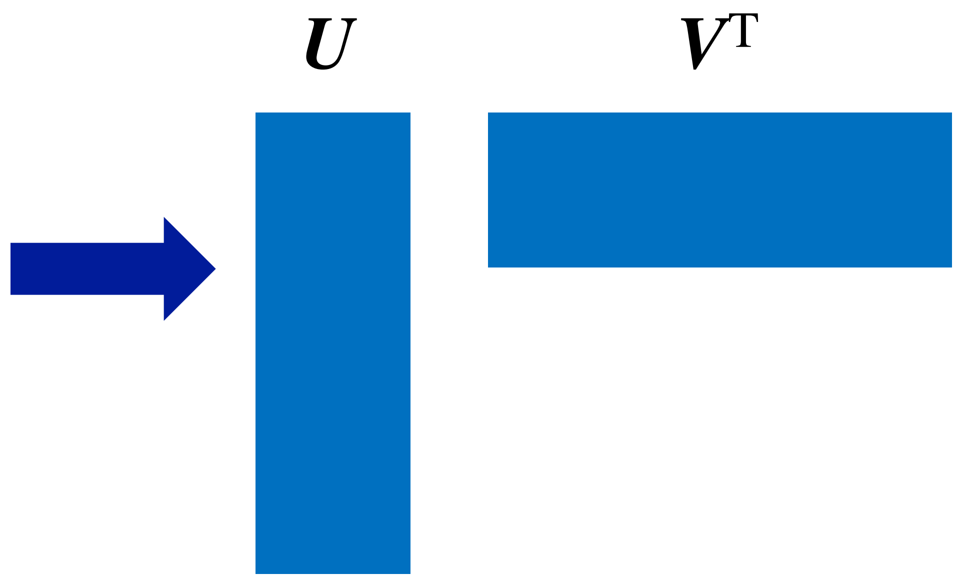
Tang et al. **Social Recommendation: A Review.** *Social Network Analysis and Mining*, 2013. Springer.

Matrix Factorization (MF) based CF

- Low-rank MF on the user-item rating matrix ***R***
- User preference vector ***U***
- Item characteristic vector ***V***

R

	v_1	v_2	v_3	v_4	v_5
u_1	5	?	2	?	?
u_2	4	4	5	?	?
u_3	?	?	4	4	1
u_4	?	?	?	?	3
u_5	?	?	1	?	?



Koren. **Factorizatoin Meets the Neighborhood: A Multifaceted Collaborative Filtering Model.** *KDD*, **2008**.

Matrix Factorization (MF) based CF

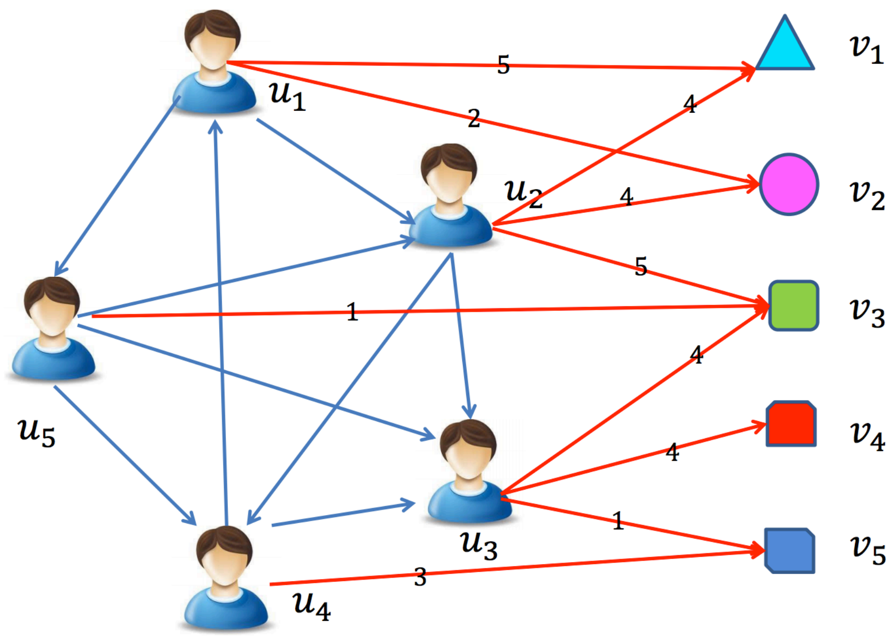
- Low-rank MF on the user-item rating matrix ***R***
- User preference vector ***U***
- Item characteristic vector ***V***
- Observed weight matrix ***W***

$$\min_{\mathbf{U},\mathbf{V}} \sum_{i=1}^{n}\sum_{j=1}^{m} \boxed{\mathbf{W}_{ij}} (\mathbf{R}_{ij} - \mathbf{U}_i\mathbf{V}_j^{\top})^2 + \boxed{\alpha(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2)}$$

avoid over-fitting,
controlled by the parameter

Social Recommendation

Social relations



	u_1	u_2	u_3	u_4	u_5
u_1	0	1	0	0	1
u_2	0	0	1	1	0
u_3	0	0	0	0	0
u_4	1	0	1	0	0
u_5	0	1	1	1	0

	v_1	v_2	v_3	v_4	v_5
u_1	5	?	2	?	?
u_2	4	4	5	?	?
u_3	?	?	4	4	1
u_4	?	?	?	?	3
u_5	?	?	1	?	?

Memory based Social Recommender

- TidalTrust

rating
(user i, item m)

$$r_{sm} = \frac{\sum_{i \in S} t_{si} r_{im}}{\sum_{i \in S} t_{si}}$$

rating
(user s, item m)

trust from social relation
(user s, user i)

Golbeck. **Personalizing applications through integration of inferred trust values in semantic web-based social networks.** *Semantic Network Analysis Workshop*, 2005.

Memory based Social Recommender

- MoleTrust

rating (user u, item i)

average rating (user u)

average rating (user a)

$$p_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^{k} w_{a,u}(r_{u,i} - \bar{r}_u)}{\sum_{u=1}^{k} w_{a,u}}$$

predicted rating (user a, item i)

trust from social relation (user a, user u)

Massa et al. **Trust-aware recommender systems.** *RecSys*, 2007.

Memory based Social Recommender

- **TrustWalker**

similarity measure
(item i, item j)

$$P(Y_{u,i} = j) = \frac{sim(i,j)}{\sum_{l \in RI_u} sim(i,l)}$$

probability of
user u's random walk
from item i to item j

Pearson correlation
of (item i, item j)

$$sim(i,j) = \frac{1}{1 + e^{-\frac{|UC_{i,j}|}{2}}} \times corr(i,j)$$

common user set
of (item i, item j)

Jamali et al. **TrustWalker: A Random Walk Model for Combining Trust-based and Item-based Recommendation.** *KDD*, 2009.

Model based Social Recommender

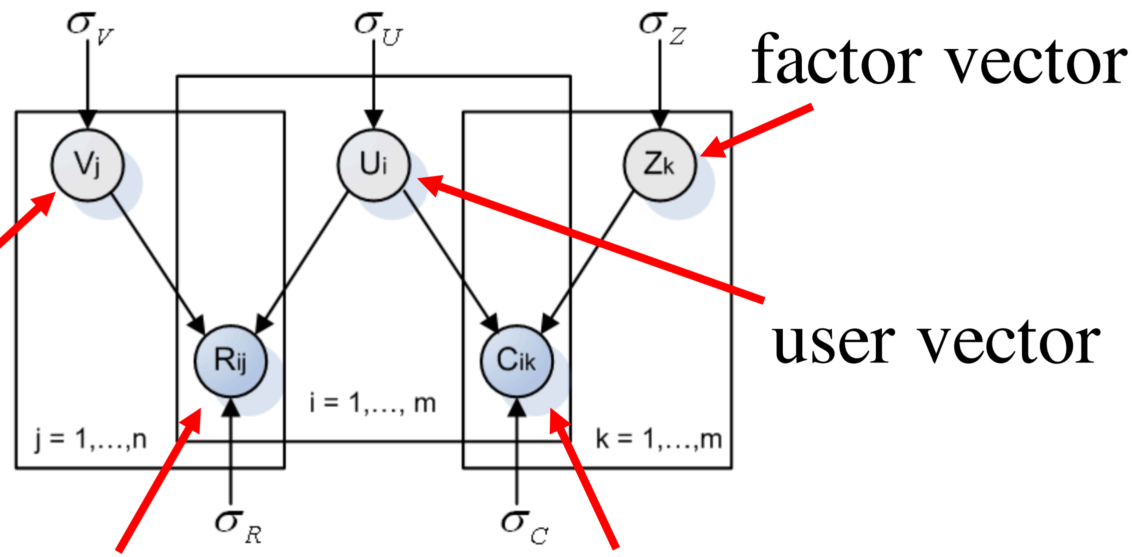
- Optimization methods such as gradient based methods can be applied to find a well-worked optimal solution.
- MF has a nice probabilistic interpretation with Gaussian noise.
- MF is very flexible and allows us to include prior knowledge.

$$
\begin{aligned}
& \textit{Social Recommendation CF} \\
= \; & \textit{Basic CF} + \textit{Social Information Model}
\end{aligned}
$$

Tang et al. **Social Recommendation: A Review.** *Social Network Analysis and Mining*, 2013. Springer.

Model based Social Recommender

- **SoRec**



***R*:** user-item rating matrix

	v_1	v_2	v_3	v_4	v_5
u_1	5	?	2	?	?
u_2	4	4	5	?	?
u_3	?	?	4	4	1
u_4	?	?	?	?	3
u_5	?	?	1	?	?

***C*:** user-user social matrix

	u_1	u_2	u_3	u_4	u_5
u_1	0	1	0	0	1
u_2	0	0	1	1	0
u_3	0	0	0	0	0
u_4	1	0	1	0	0
u_5	0	1	1	1	0

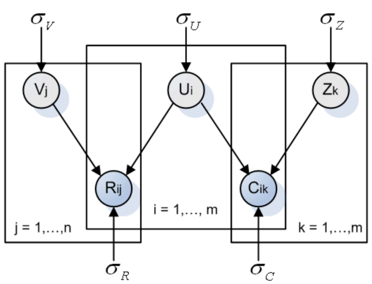
Ma et al. **Social Recommendation Using Probabilistic Matrix Factorization.** *CIKM*, 2008.
Improving Recommender Systems by Incorporating Social Contextual Information. *TIS*, 2011.

Model based Social Recommender

- SoRec

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

R



Gaussian distribution

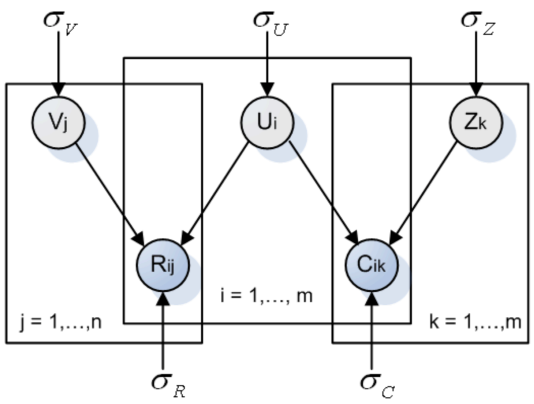
Logistic function

Observed

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

Model based Social Recommender

- **SoRec**



behavioral term

social term

$$\mathcal{L}(R,C,U,V,Z) =$$
$$\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}^{R}\underline{(r_{ij}-g(U_i^T V_j))^2}+\frac{\lambda_C}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} I_{ik}^{C}\underline{(c_{ik}^{*}-g(U_i^T Z_k))^2}$$
$$\underline{+\frac{\lambda_U}{2}\|U\|_F^2 + \frac{\lambda_V}{2}\|V\|_F^2 + \frac{\lambda_Z}{2}\|Z\|_F^2}, \qquad (9)$$

regularization terms

Model based Social Recommender

❑SoRec

Gradient Descent Methods

deviate of Logistic function

$$\frac{\partial \mathcal{L}}{\partial U_i} = \sum_{j=1}^{n} I_{ij}^{R} \underline{g'(U_i^T V_j)}(g(U_i^T V_j) - r_{ij})V_j$$

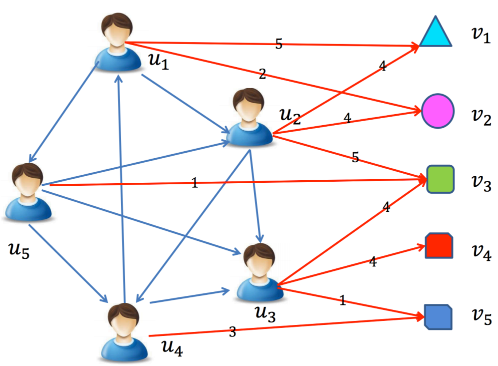
$$+ \lambda_C \sum_{j=1}^{m} I_{ik}^{C} \underline{g'(U_i^T Z_k)}(g(U_i^T Z_k) - c_{ik}^{*})Z_k + \lambda_U U_i,$$

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

$$\frac{\partial \mathcal{L}}{\partial Z_k} = \lambda_C \sum_{i=1}^{m} I_{ik}^{C} \underline{g'(U_i^T Z_k)}(g(U_i^T Z_k) - c_{ik}^{*})U_i + \lambda_Z Z_k, (10)$$

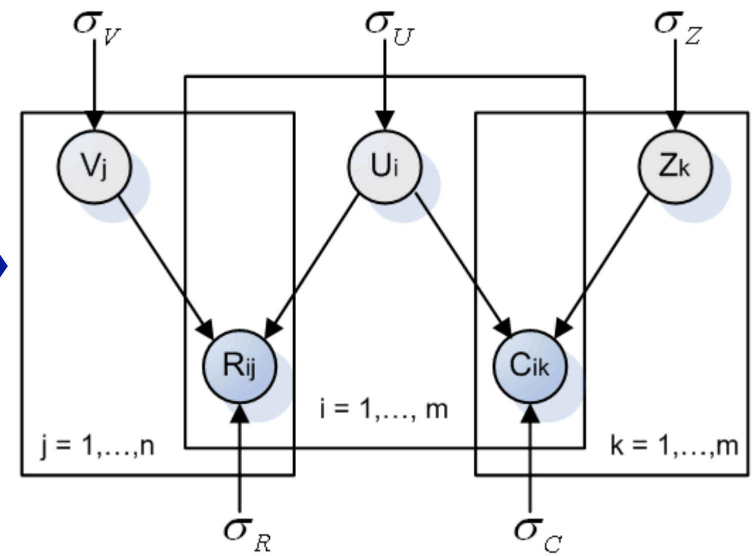
Model based Social Recommender

- SoRec



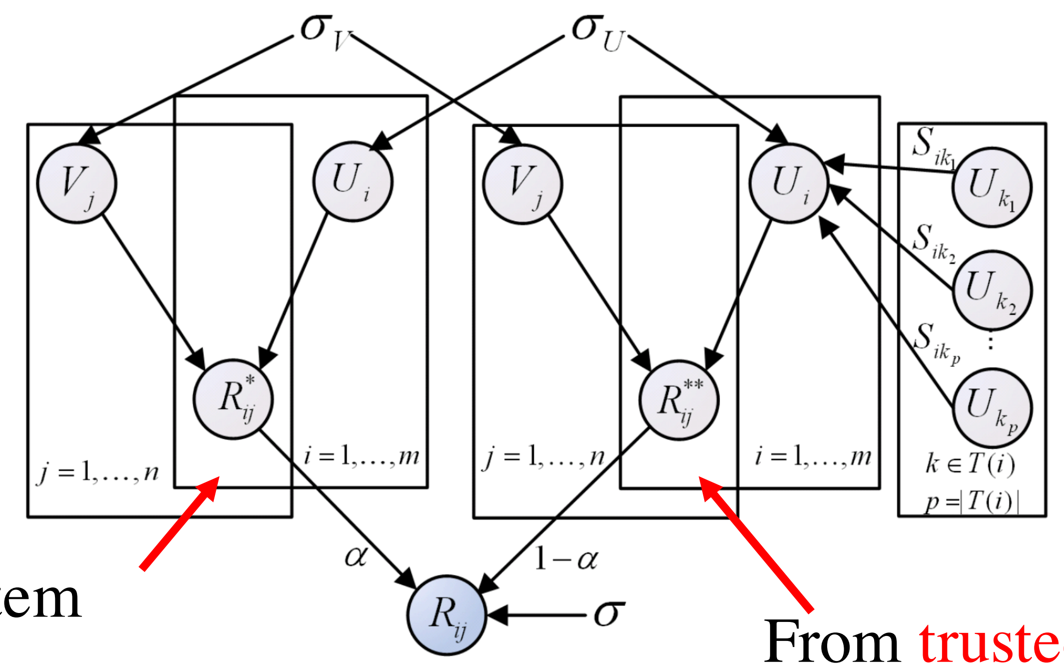
	u_1	u_2	u_3	u_4	u_5
u_1	0	1	0	0	1
u_2	0	0	1	1	0
u_3	0	0	0	0	0
u_4	1	0	1	0	0
u_5	0	1	1	1	0

	v_1	v_2	v_3	v_4	v_5
u_1	5	?	2	?	?
u_2	4	4	5	?	?
u_3	?	?	4	4	1
u_4	?	?	?	?	3
u_5	?	?	1	?	?



Model based Social Recommender

- Replacing social with trust
- **"Social Trust" Ensemble** for Epinion data



From user-item matrix

Operation "plus"

From trusted friends

Ma et al. **Learning to Recommend with Social Trust Ensemble.** *SIGIR*, 2009.
Learning to Recommend with Explicit and Implicit Social Relations. *TIST*, 2011.

Model based Social Recommender

- "Social Trust" Ensemble

From user-item matrix

From trusted friends

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

Model based Social Recommender

- "Social Trust" Ensemble

Gradient Descent Methods

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial U_i} = & \alpha \sum_{j=1}^{n} I_{ij}^R g'(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) V_j \\
& \times (g(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) - R_{ij}) \\
& + (1-\alpha) \sum_{p \in \mathcal{B}(i)} \sum_{j=1}^{n} I_{pj}^R g'(\alpha U_p^T V_j + (1-\alpha) \sum_{k \in \mathcal{T}(p)} S_{pk} U_k^T V_j) \\
& \times (g(\alpha U_p^T V_j + (1-\alpha) \sum_{k \in \mathcal{T}(p)} S_{pk} U_k^T V_j) - R_{pj}) S_{pi} V_j + \lambda_U U_i, \\
\frac{\partial \mathcal{L}}{\partial V_j} = & \sum_{i=1}^{m} I_{ij}^R g'(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) \\
& \times (g(\alpha U_i^T V_j + (1-\alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T V_j) - R_{ij}) \\
& \times (\alpha U_i + (1-\alpha) \sum_{k \in \mathcal{T}(i)} S_{ik} U_k^T) + \lambda_V V_j, \qquad (14)
\end{aligned}$$

Model based Social Recommender

- **SoReg**

Average-based regularization:

Regularize with the average of friends' tastes

$$\min_{U,V} \mathcal{L}_1(R,U,V) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}(R_{ij} - U_i^T V_j)^2$$
$$\boxed{+ \frac{\alpha}{2}\sum_{i=1}^{m} \|U_i - \frac{\sum_{f\in\mathcal{F}^+(i)} Sim(i,f)\times U_f}{\sum_{f\in\mathcal{F}^+(i)} Sim(i,f)}\|_F^2,}$$
$$+ \frac{\lambda_1}{2}\|U\|_F^2 + \frac{\lambda_2}{2}\|V\|_F^2. \quad (8)$$

Information loss: Friends may have diverse tastes!!!

Ma et al. **Recommender Systems with Social Regularization.** *WSDM*, 2011.

Model based Social Recommender

- SoReg

Individual-based regularization:
Regularize with friends individually

$$\min_{U,V} \mathcal{L}_2(R,U,V) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}(R_{ij} - U_i^T V_j)^2$$
$$+ \boxed{\frac{\beta}{2}\sum_{i=1}^{m}\sum_{f\in\mathcal{F}^+(i)} Sim(i,f)\|U_i - U_f\|_F^2}$$
$$+ \lambda_1\|U\|_F^2 + \lambda_2\|V\|_F^2. \qquad (11)$$

Related Work

	Behavior	Content	Social	Trust
Collaborative filtering (CF) [Herlocker *et al*. TOIS; Koren KDD]	✔			
Content-based filtering with CF [Balabanovic *et al*.; Liu *et al*. CIKM;]	✔	✔		
SoRec [Ma *et al*. CIKM, TIS] SoReg [Ma *et al*. WSDM]	✔		✔	
Trust-based methods [Massa *et al*. RecSys; Jamali *et al*. KDD; Ma *et al*. SIGIR, TIST]	✔			✔

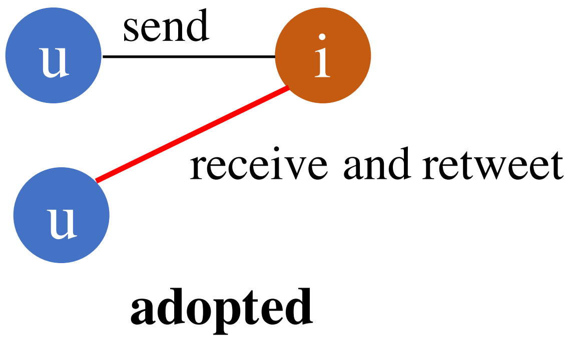
- **Q:** What are the **factors** of users' decisions on retweeting? Can we **observe** them from the data? How to **integrate** the information for accurate prediction?

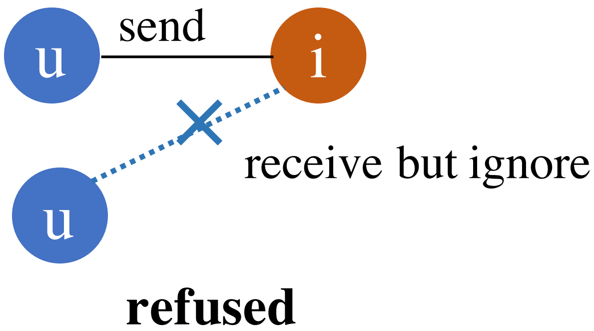
Observation: Social Contextual Factors

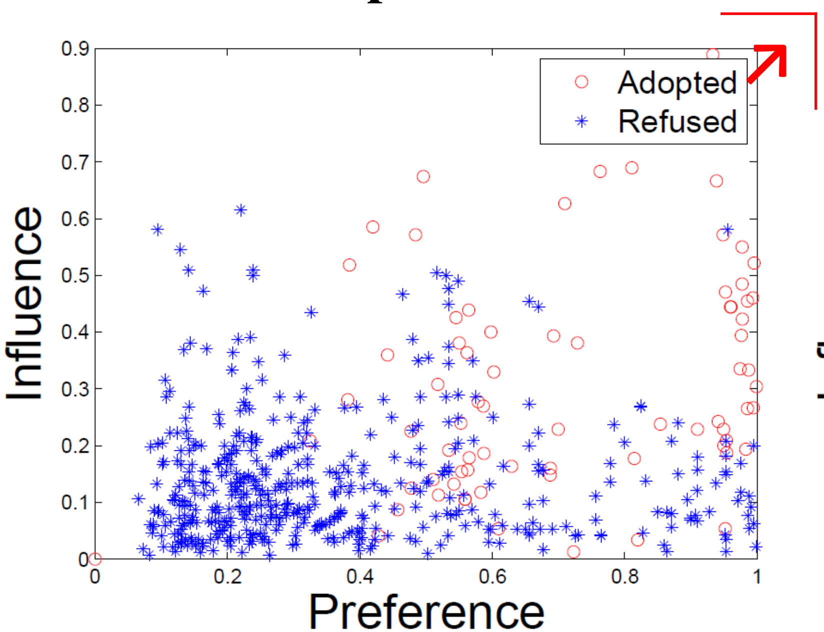
- Will Michelle Obama share this message?
- Please list your reasons.

Jiang et al. **Social Contextual Recommendation.** *CIKM*, 2012.
Social Recommendation with Contextual Information. *TKDE*, 2014.

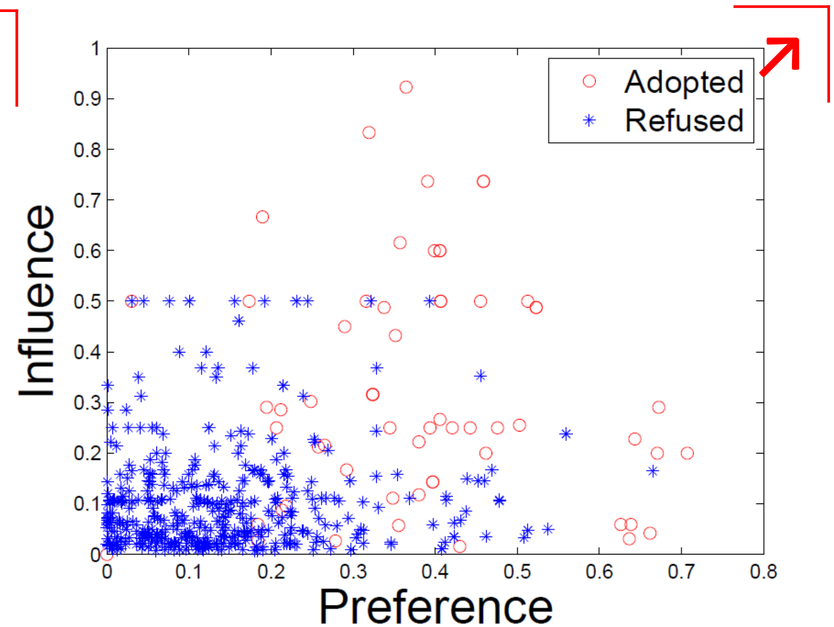
Observation: Social Contextual Factors





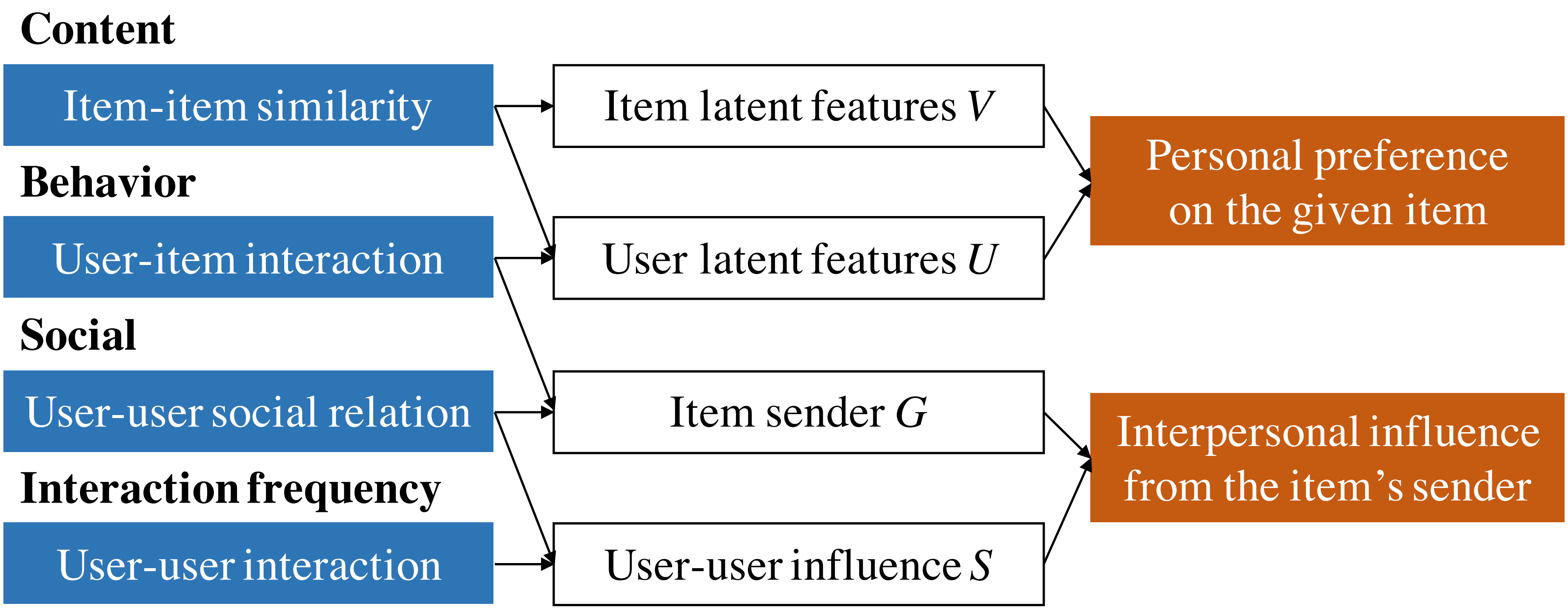


China's Facebook: Renren

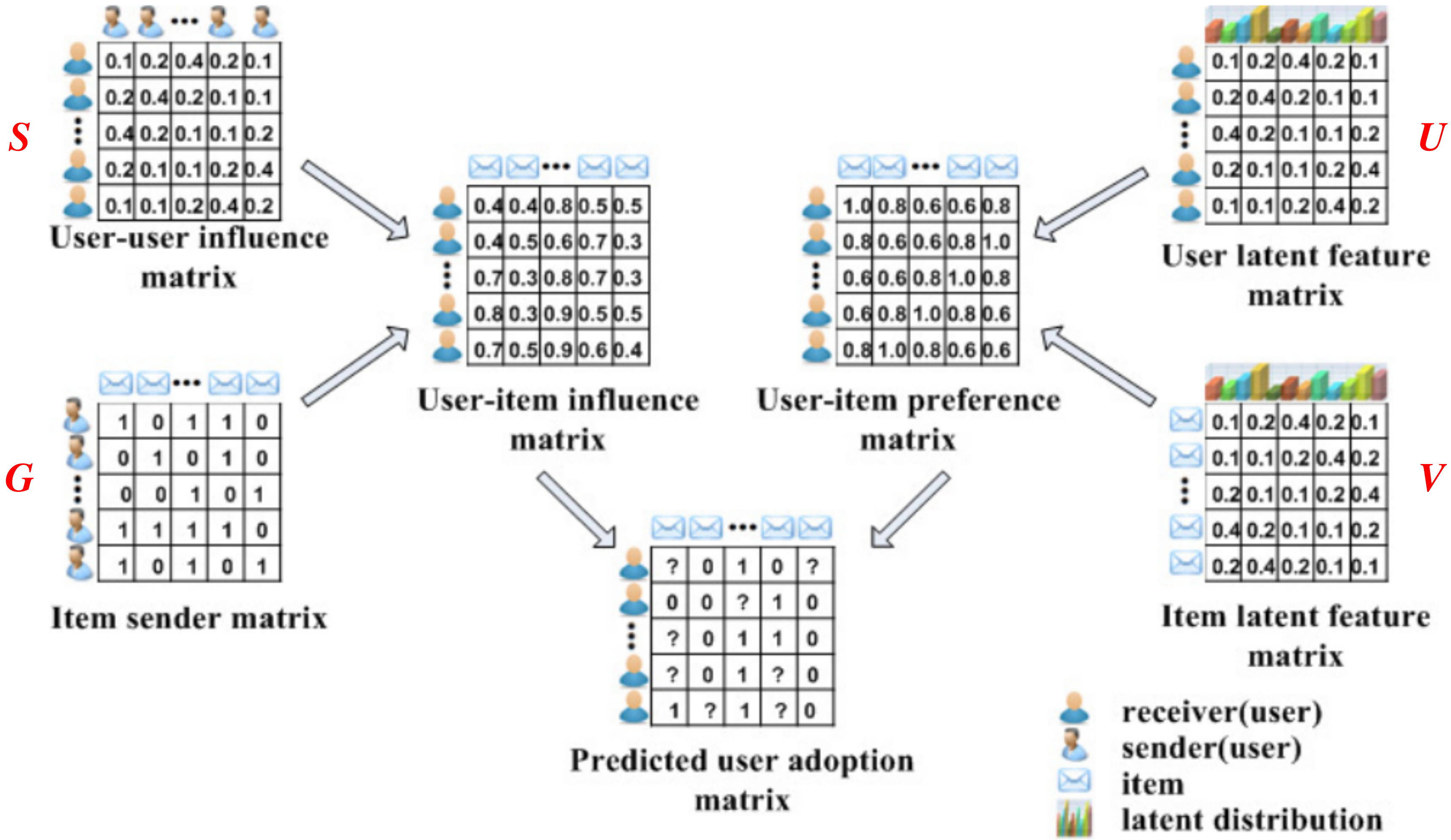


China's Twitter: Tencent Weibo

Representation: From Contextual Information to Contextual Factors



Model: ContextMF



Model: ContextMF

behavior influence preference

$$P(\mathbf{R}|\mathbf{S},\mathbf{U},\mathbf{V},\sigma_R^2) = \prod_{i=1}^{M}\prod_{j=1}^{N}\mathcal{N}(\mathbf{R}_{ij}|\mathbf{S}_i\mathbf{G}_j^\top \odot \mathbf{U}_i^\top \mathbf{V}_j, \sigma_R^2)$$

behavior

interaction frequency/trust

item content

$$\mathcal{J} = ||\mathbf{R} - \mathbf{S}\mathbf{G}^\top \odot \mathbf{U}^\top\mathbf{V}||_F^2 + \alpha||\mathbf{W} - \mathbf{U}^\top\mathbf{U}||_F^2$$
$$+\beta||\mathbf{C} - \mathbf{V}^\top\mathbf{V}||_F^2 + \gamma||\mathbf{S} - \mathbf{F}||_F^2$$
$$+\delta||\mathbf{S}||_F^2 + \eta||\mathbf{U}||_F^2 + \lambda||\mathbf{V}||_F^2$$

social relation

Model: ContextMF

- Gradient descent method

$$
\begin{aligned}
\frac{\partial \mathcal{J}}{\partial \mathbf{S}} &= 2\Big(-\mathbf{R}(\mathbf{G} \odot \mathbf{V}^\top \mathbf{U}) + (\mathbf{S}\mathbf{G}^\top \odot \mathbf{U}^\top \mathbf{V})\mathbf{G} \\
&\quad + \gamma(\mathbf{S} - \mathbf{F}) + \delta \mathbf{S}\Big) \\
\frac{\partial \mathcal{J}}{\partial \mathbf{U}} &= 2\Big(-\mathbf{V}\mathbf{R}^\top + \mathbf{V}(\mathbf{G}\mathbf{S}^\top \odot \mathbf{V}^\top \mathbf{U}) - 2\alpha \mathbf{U}\mathbf{W} \\
&\quad + 2\alpha \mathbf{U}\mathbf{U}^\top \mathbf{U} + \eta \mathbf{U}\Big) \\
\frac{\partial \mathcal{J}}{\partial \mathbf{V}} &= 2\Big(-\mathbf{U}\mathbf{R} + \mathbf{U}(\mathbf{S}\mathbf{G}^\top \odot \mathbf{U}^\top \mathbf{V}) - 2\beta \mathbf{V}\mathbf{C} \\
&\quad + 2\beta \mathbf{V}\mathbf{V}^\top \mathbf{V} + \lambda \mathbf{V}\Big)
\end{aligned}
$$

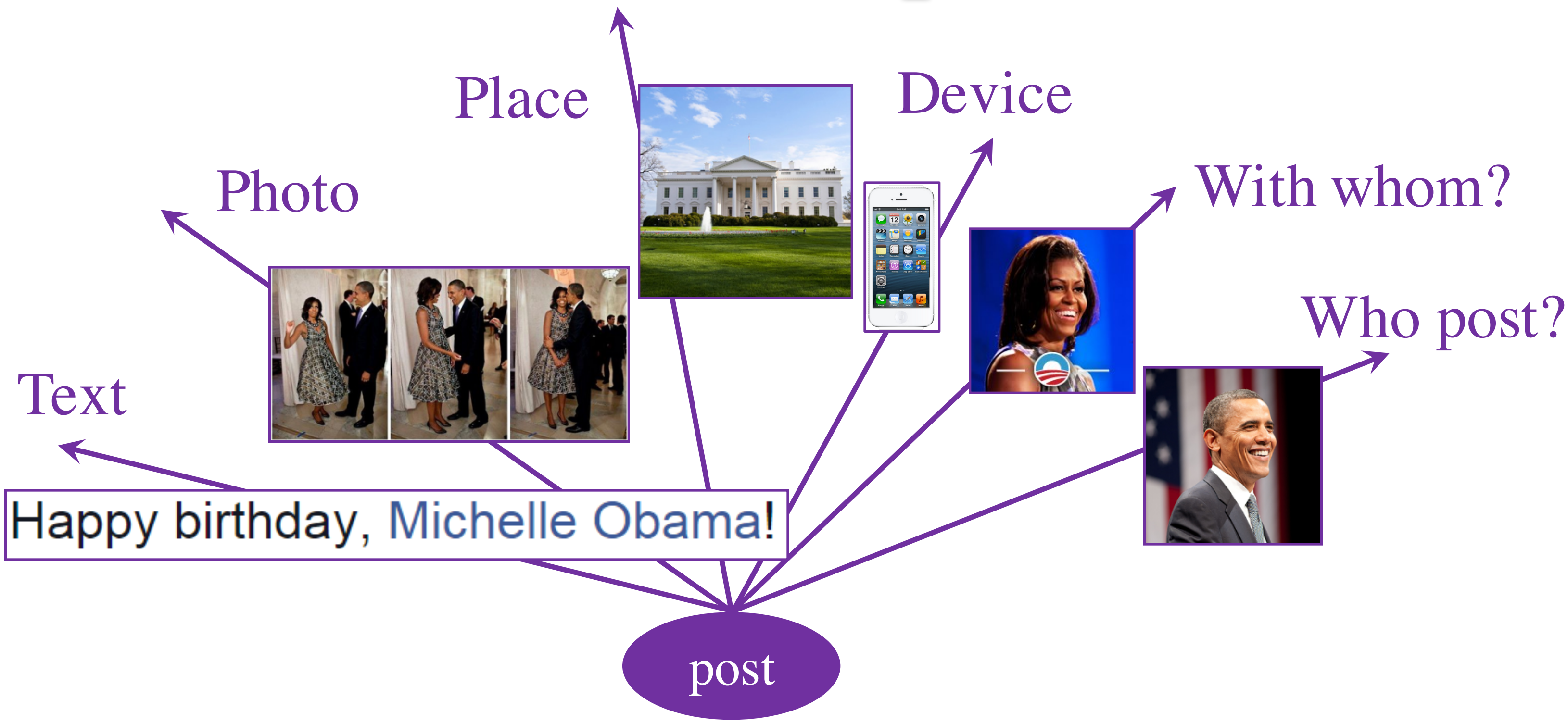
Experimental Results

Method	MAE	RMSE	$\hat{\tau}$	$\hat{\rho}$
Renren Dataset				
Content-based [1]	0.3842	0.4769	0.5409	0.5404
Item CF [25]	0.3601	0.4513	0.5896	0.5988
FeedbackTrust [22]	0.3764	0.4684	0.5433	0.5469
Influence-based [9]	0.3859	0.4686	0.5394	0.5446
SoRec [19]	0.3276	0.4127	0.6168	0.6204
SoReg [20]	0.2985	0.3537	0.7086	0.7140
Influence MF	0.3102	0.3771	0.6861	0.7006
Preference MF	0.3032	0.3762	0.6937	0.7036
Context MF	**0.2416**	**0.3086**	**0.7782**	**0.7896**
Tencent Weibo Dataset				
Content-based [1]	0.2576	0.3643	0.7728	0.7777
Item CF [25]	0.2375	0.3372	0.7867	0.8049
FeedbackTrust [22]	0.2830	0.3887	0.7094	0.7115
Influence-based [9]	0.2651	0.3813	0.7163	0.7275
SoRec [19]	0.2256	0.3325	0.7973	0.8064
SoReg [20]	0.1997	0.2962	0.8390	0.8423
Influence MF	0.2183	0.3206	0.8179	0.8258
Preference MF	0.2111	0.3088	0.8384	0.8453
Context MF	**0.1514**	**0.2348**	**0.8570**	**0.8685**

vs. SoReg [TIST'11]	Renren	Tencent Weibo
MAE	↓19.1%	↓24.2%
RMSE	↓12.8%	↓20.7%
Kendall's	↑9.82%	↑2.1%
Spearman's	↑10.6%	↑3.1%

- **Deployed** in Weibo News Feed. Improved conversion rate from 5.78% to 8.27% (relatively **43%**).
- #citations = **149**

Observation: Spatial Context



Jan. 18
Birthday party
@ White house

Jiang et al. **Flexible Evolutionary Multi-faceted Analysis for Dynamic Behavioral Pattern Discovery.** *KDD*, 2014.

Observation: Temporal Context

Photo

Text

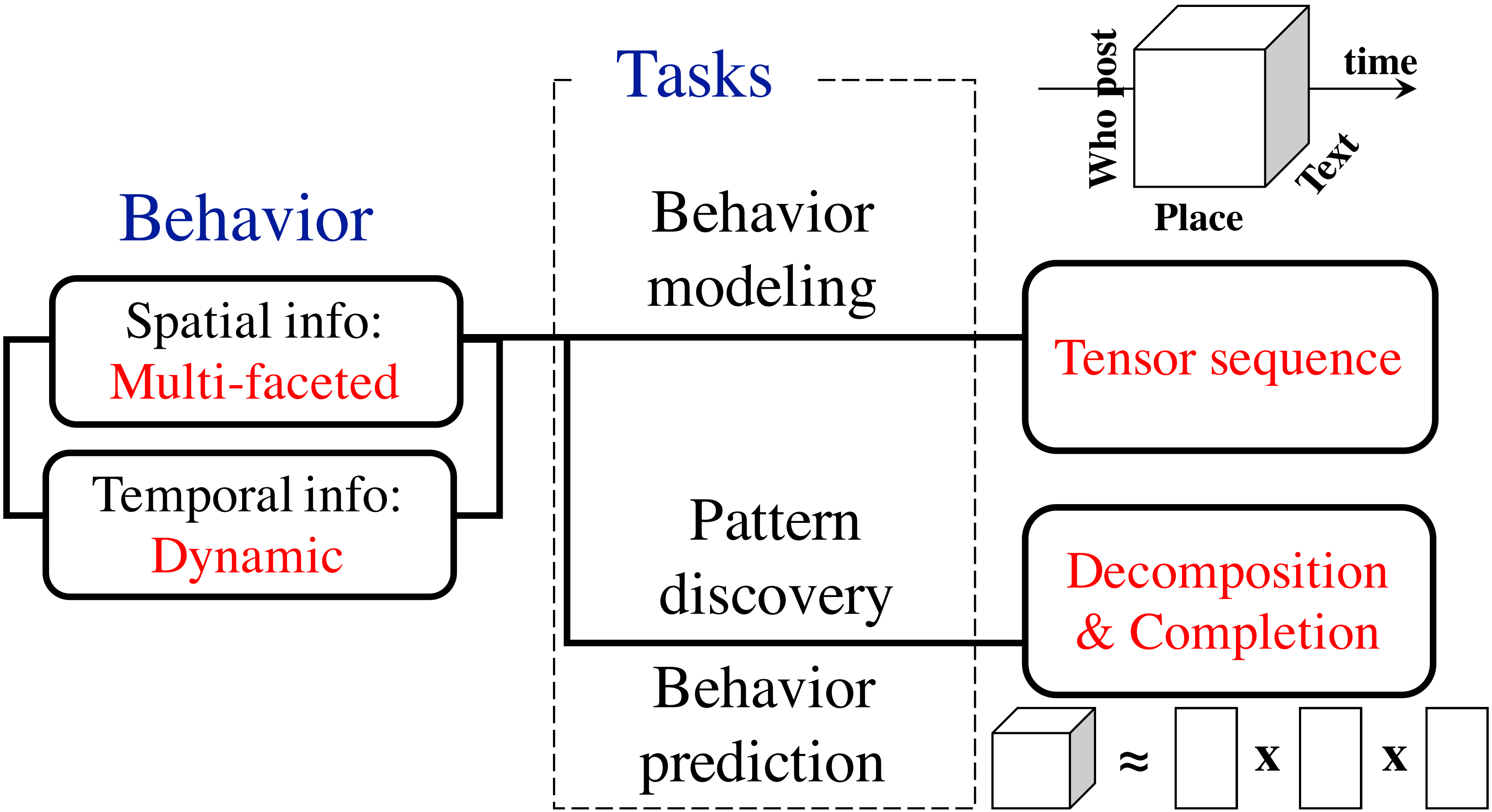
Happy birthday, Michelle Obama!

post → post → time

Jan. 18
Birthday party
@ White house

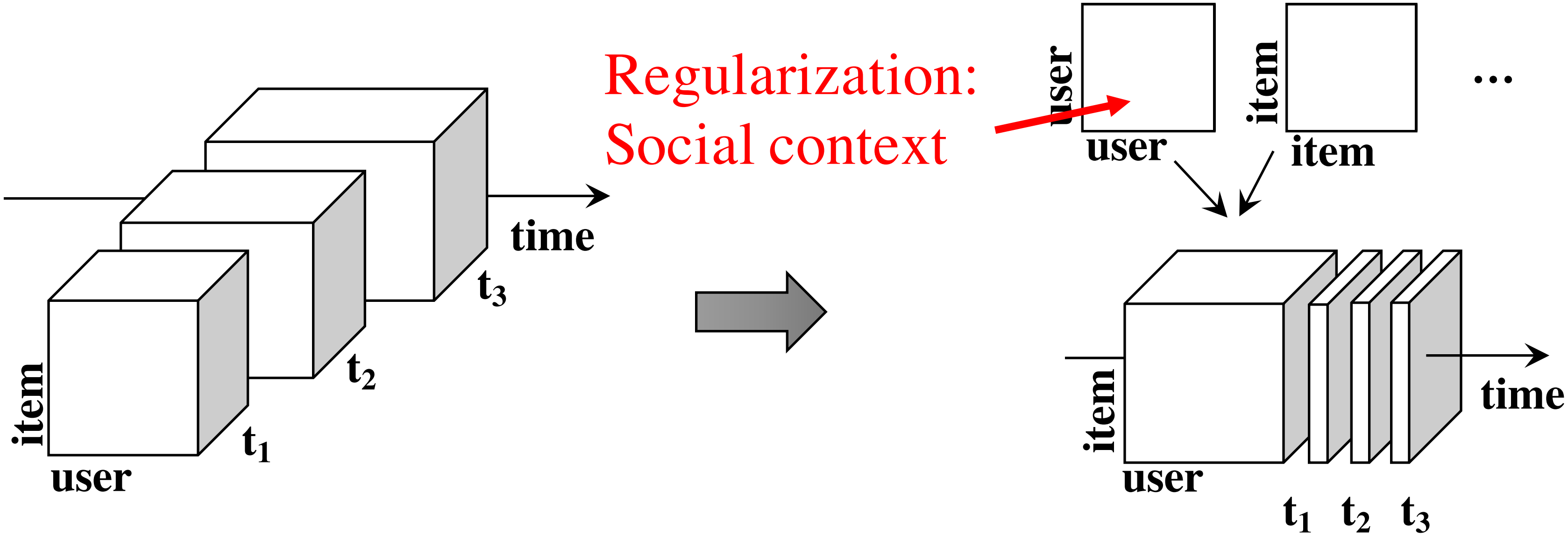
March 17
St. Patrick's Day
@ Bar

Representation: Tensor Sequence



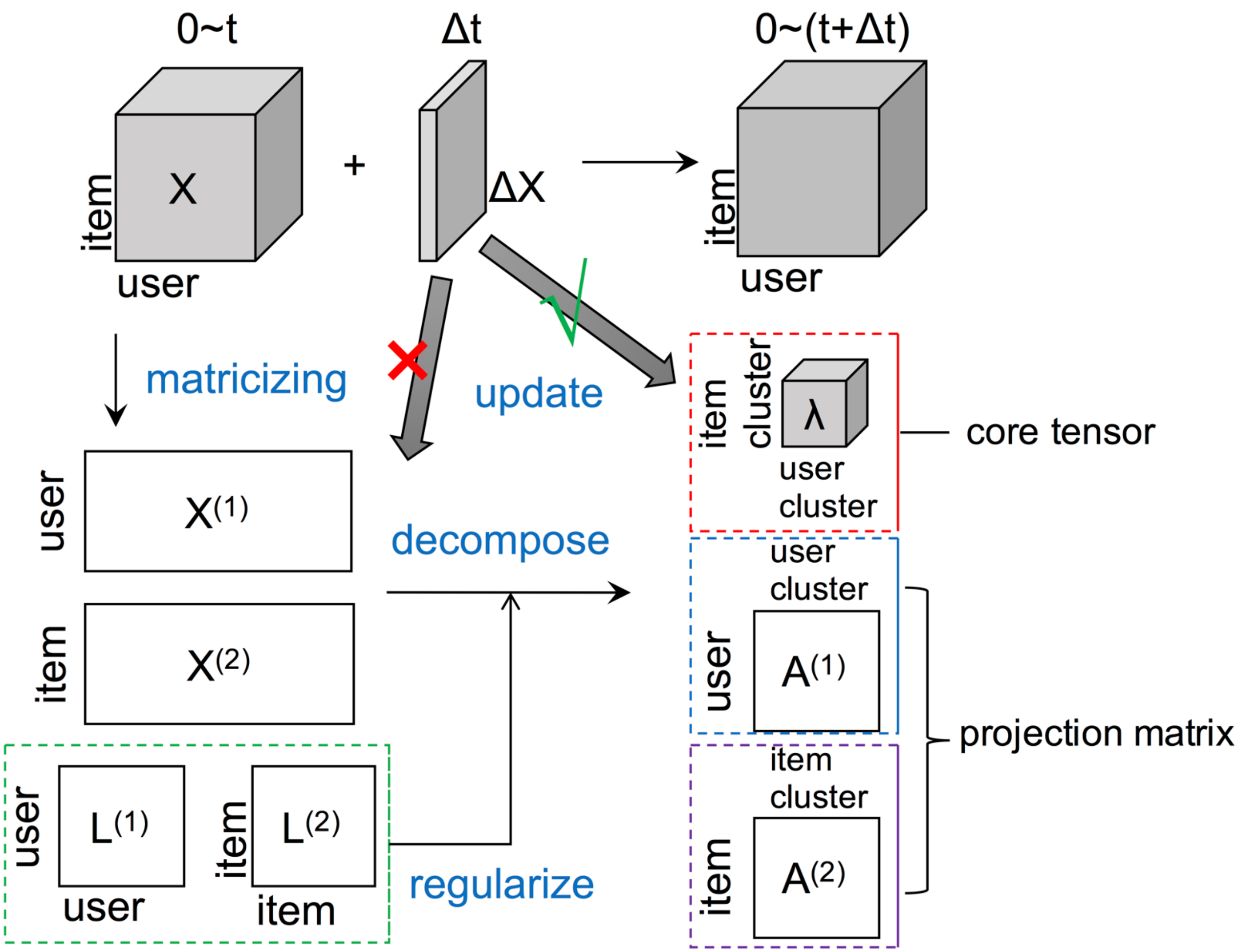
Challenges: Sparsity and Complexity

- Addressing sparsity: *Flexible regularization with auxiliary data*
- Addressing high complexity: *Incremental updates for projection matrix*



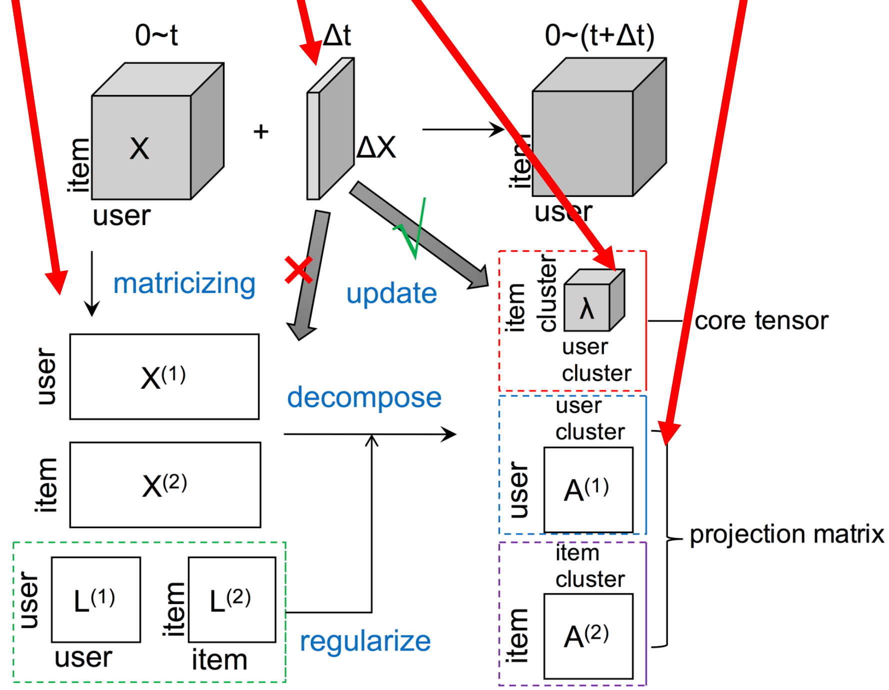
Model: FEMA

Flexible Evolutionary Multi-faceted Analysis



Tensor Perturbation Theory

$$[(\mathbf{X}^{(\mathbf{m})} + \Delta\mathbf{X}^{(\mathbf{m})})(\mathbf{X}^{(\mathbf{m})} + \Delta\mathbf{X}^{(\mathbf{m})})^{\top} + \mu^{(m)}\mathbf{L}^{(\mathbf{m})}] \cdot(\mathbf{a}_{\mathbf{i}}^{(\mathbf{m})} + \Delta\mathbf{a}_{\mathbf{i}}^{(\mathbf{m})}) = (\lambda_i^{(m)} + \Delta\lambda_i^{(m)})(\mathbf{a}_{\mathbf{i}}^{(\mathbf{m})} + \Delta\mathbf{a}_{\mathbf{i}}^{(\mathbf{m})})$$



Algorithm: FEMA

Approximation

Require: $\mathcal{X}_\mathbf{t}, \Delta\mathcal{X}_\mathbf{t}, \mathbf{A}_\mathbf{t}^{(\mathbf{m})}|_{m=1}^{M}, \lambda_\mathbf{t}^{(\mathbf{m})}|_{m=1}^{M}$

for $m = 1, \dots, M$ **do**

  for $i = 1, \dots, r^{(m)}$ **do**

    Compute $\Delta\lambda_{t,i}^{(m)}$ using

$$\boxed{\Delta\lambda_i^{(m)} = \mathbf{a}_\mathbf{i}^{(\mathbf{m})\top}(\mathbf{X}^{(\mathbf{m})}\Delta\mathbf{X}^{(\mathbf{m})\top} + \Delta\mathbf{X}^{(\mathbf{m})}\mathbf{X}^{(\mathbf{m})\top})\mathbf{a}_\mathbf{i}^{(\mathbf{m})}}$$

    and compute

$$\lambda_{t+1,i}^{(m)} = \lambda_{t,i}^{(m)} + \Delta\lambda_{t,i}^{(m)};$$

    Compute $\Delta\mathbf{a}_{\mathbf{t},\mathbf{i}}^{(\mathbf{m})}$ using

$$\boxed{\Delta\mathbf{a}_\mathbf{i}^{(\mathbf{m})} = \sum_{j \neq i} \frac{\mathbf{a}_\mathbf{j}^{(\mathbf{m})\top}(\mathbf{X}^{(\mathbf{m})}\Delta\mathbf{X}^{(\mathbf{m})\top} + \Delta\mathbf{X}^{(\mathbf{m})}\mathbf{X}^{(\mathbf{m})\top})\mathbf{a}_\mathbf{i}^{(\mathbf{m})}}{\lambda_i^{(m)} - \lambda_j^{(m)}}\mathbf{a}_\mathbf{j}^{(\mathbf{m})}}$$

    and compute

$$\mathbf{a}_{\mathbf{t+1},\mathbf{i}}^{(\mathbf{m})} = \mathbf{a}_{\mathbf{t},\mathbf{i}}^{(\mathbf{m})} + \Delta\mathbf{a}_{\mathbf{t},\mathbf{i}}^{(\mathbf{m})} \text{ and } \mathbf{A}_{\mathbf{t+1}}^{(\mathbf{m})} = \{\mathbf{a}_{\mathbf{t+1},\mathbf{i}}^{(\mathbf{m})}\};$$

  end for

end for

$\mathcal{Y}_{\mathbf{t+1}} = (\mathcal{X}_\mathbf{t} + \Delta\mathcal{X}_\mathbf{t})\prod_{m=1}^{M} \times_{(m)} \mathbf{A}_{\mathbf{t+1}}^{(\mathbf{m})\mathbf{T}};$

return $\mathbf{A}_{\mathbf{t+1}}^{(\mathbf{m})}|_{m=1}^{M}, \lambda_{\mathbf{t+1}}^{(\mathbf{m})}|_{m=1}^{M}, \mathcal{Y}_{\mathbf{t+1}}$

Bound Guarantee

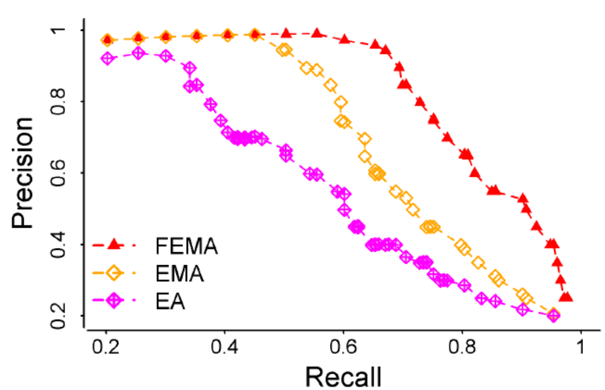
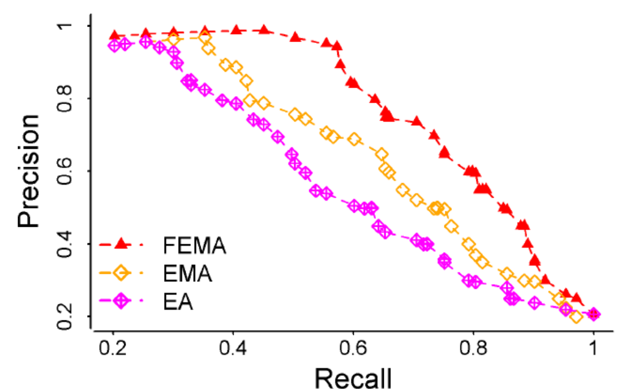
core tensor

$$|\Delta\lambda_i^{(m)}| \leq 2(\lambda^{\max}_{\mathbf{X}^{(\mathbf{m})\top}\mathbf{X}^{(\mathbf{m})}})^{\frac{1}{2}}\|\Delta\mathbf{X}^{(\mathbf{m})}\|_2$$

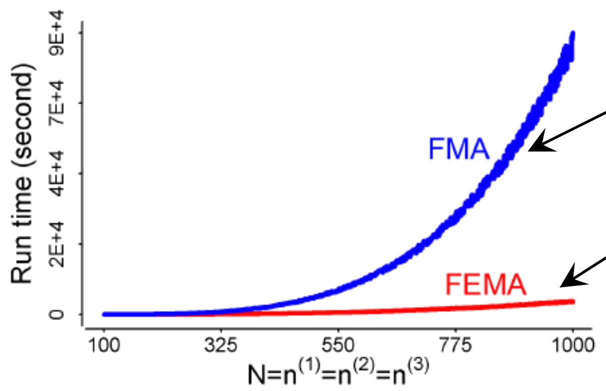
$$|\Delta\mathbf{a}_\mathbf{i}^{(\mathbf{m})}| \leq 2\|\Delta\mathbf{X}^{(\mathbf{m})}\|_2 \sum_{j \neq i} \frac{(\lambda^{\max}_{\mathbf{X}^{(\mathbf{m})\top}\mathbf{X}^{(\mathbf{m})}})^{\frac{1}{2}}}{|\lambda_i^{(m)} - \lambda_j^{(m)}|}$$

projection matrix

Results: FEMA > EMA > EA

	Microsoft Academic Search		Tencent Weibo mentions “@”	
	MAE	RMSE	MAE	RMSE
FEMA X L	**0.735**	**0.944**	**0.894**	**1.312**
EMA X	0.794	1.130	0.932	1.556
EA X	0.979	1.364	1.120	1.873
Precision vs Recall				

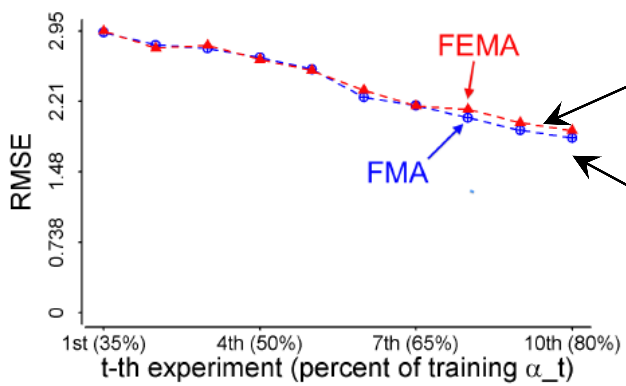
Results: Efficiency



Time vs Num. objects N

Re-decomposition:
Re-compute projection matrices

Evolutionary analysis:
Use ΔX to update λ and a



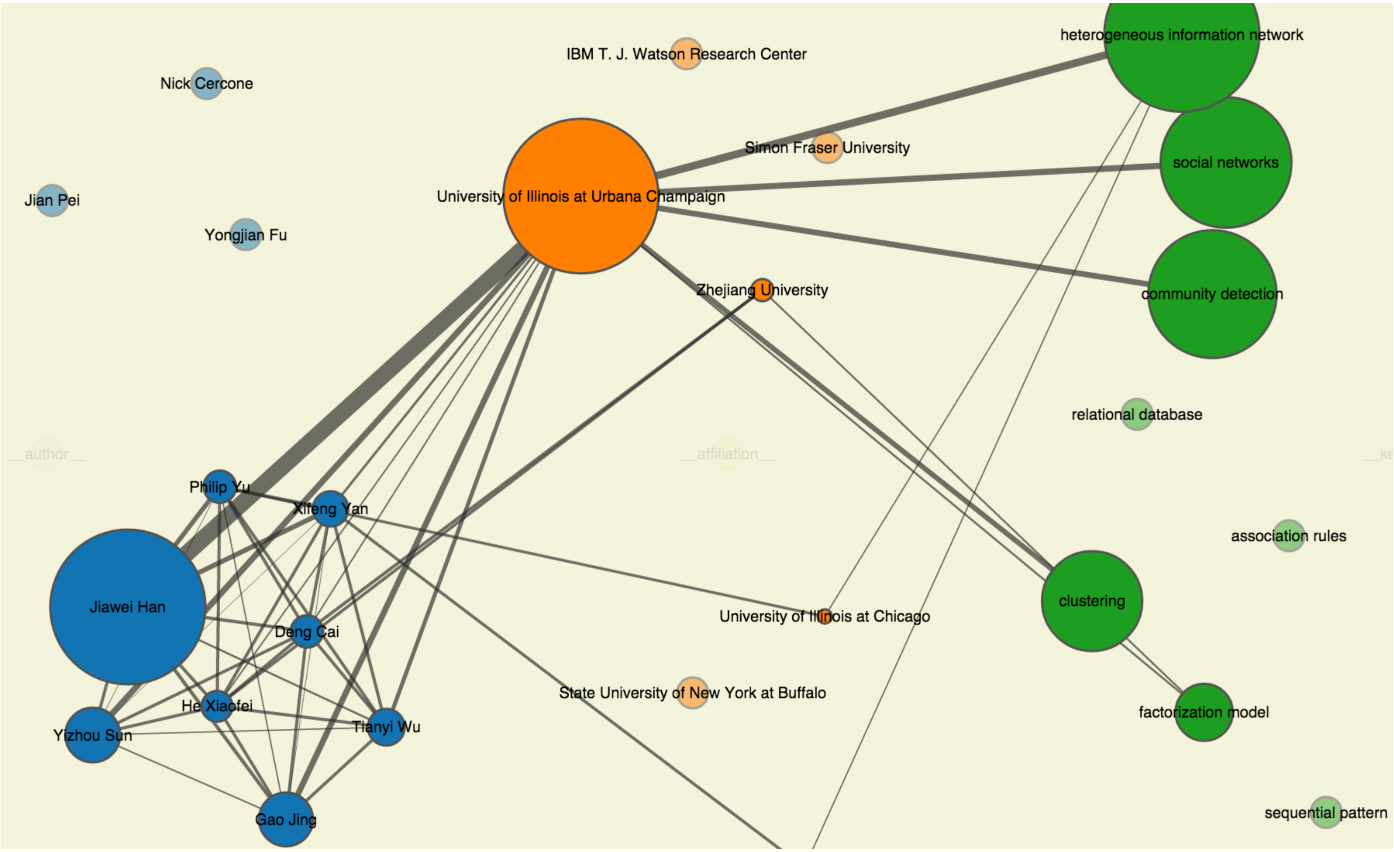
The loss is small.

Evolutionary analysis:
Use ΔX to update λ and a

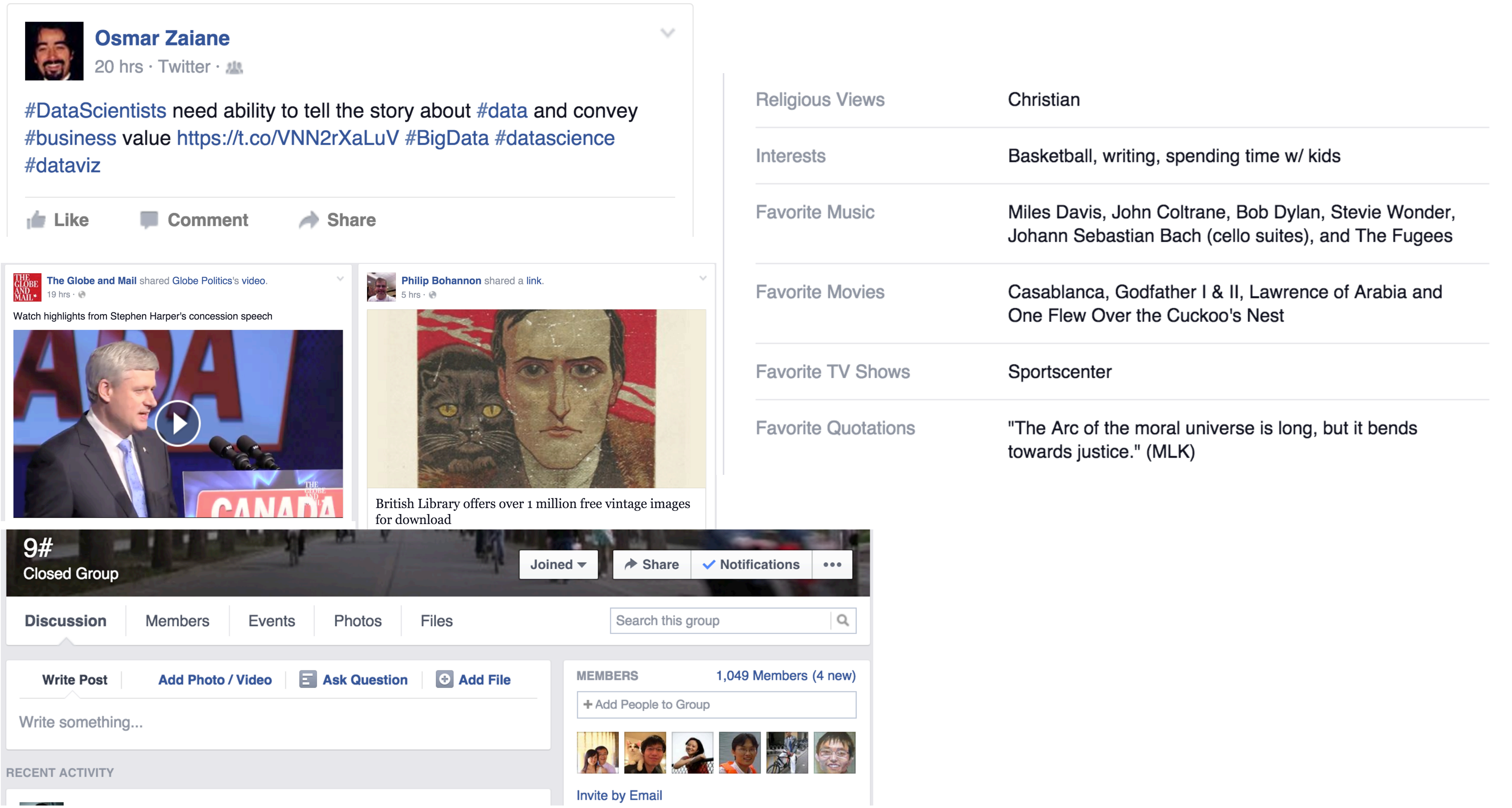
Re-decomposition:
Re-compute projection matrices

Demo: Author@Affiliation#Keyword

http://www.meng-jiang.com/demos/fema/mas/

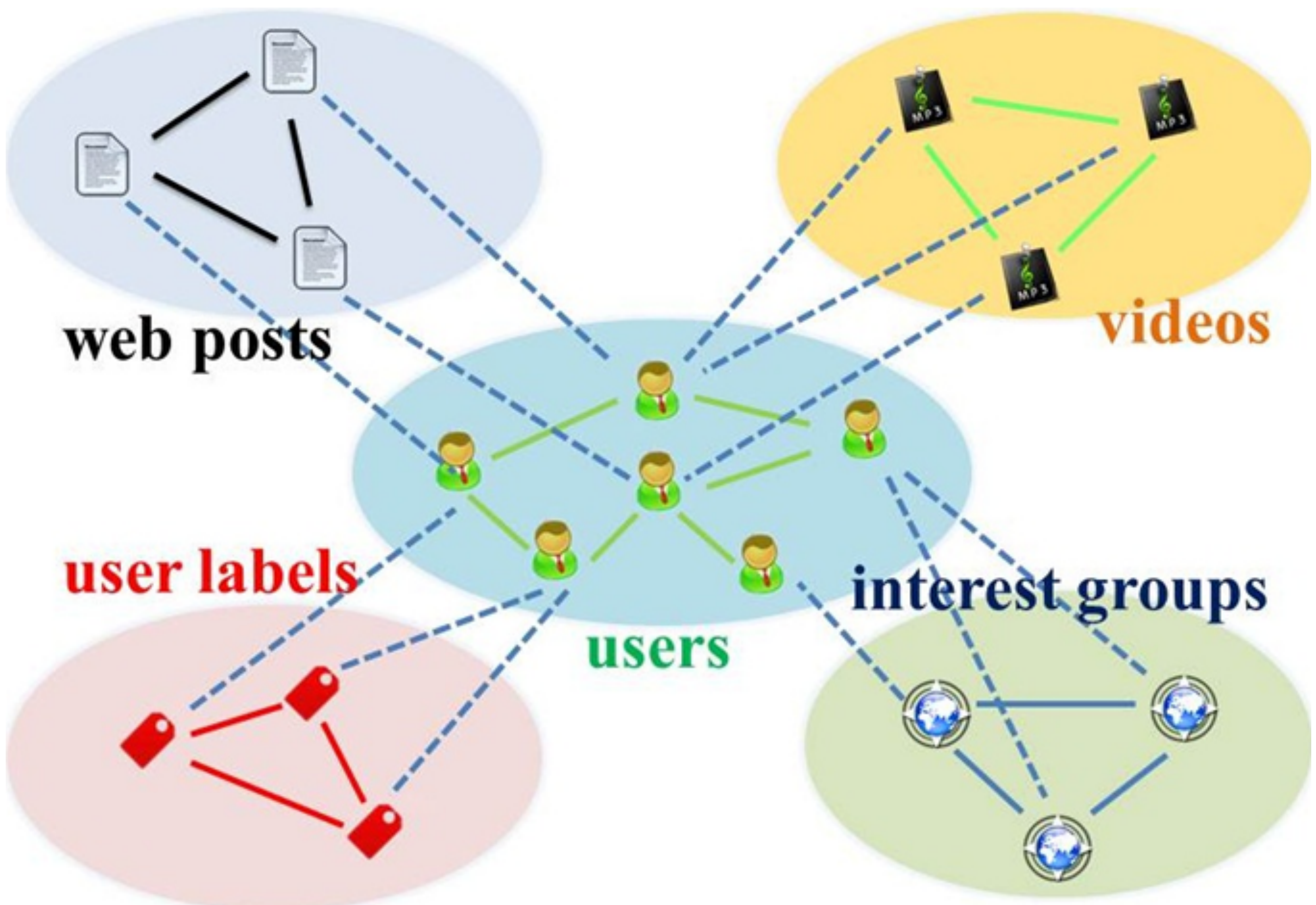


Observation: Multiple Domains

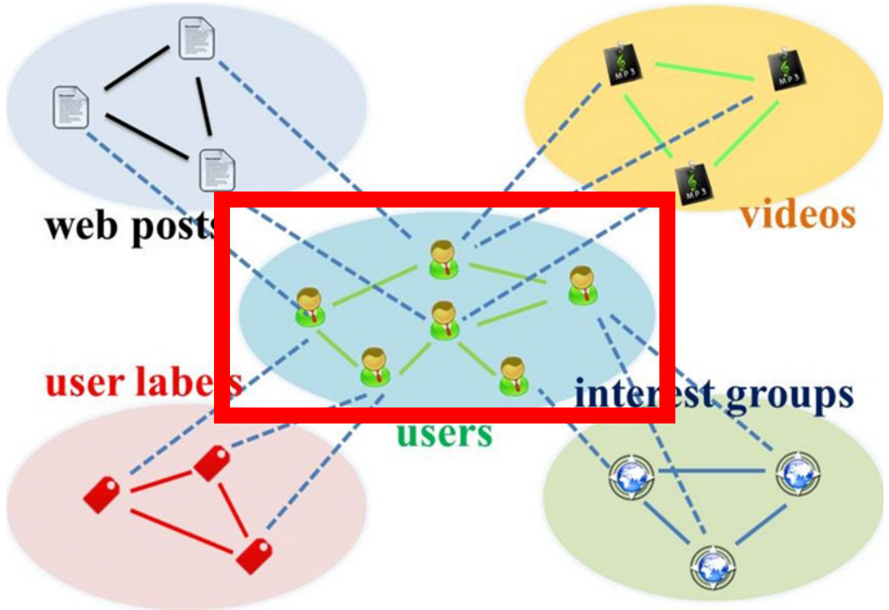


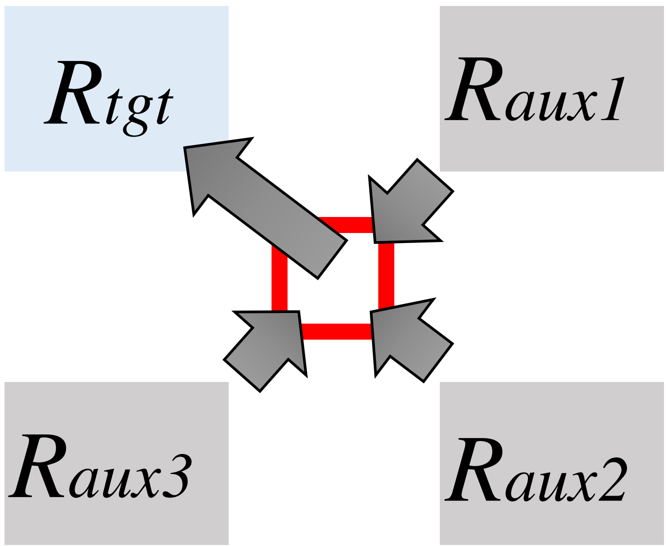
Jiang et al. **Social Recommendation across Multiple Relational Domains.** *CIKM*, 2012.
Social Recommendation with Cross-Domain Transferable Knowledge. *TKDE*, 2015.

Representation: Star-Structured Graph



Representation: Social Bridge

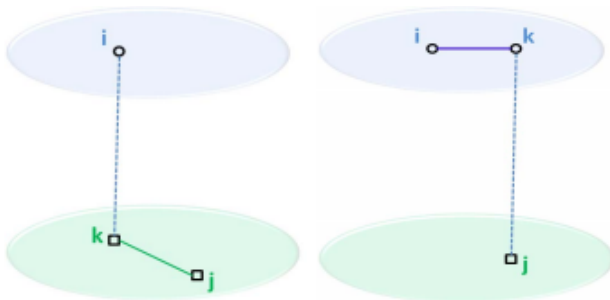




Bridge: Tie strength

Algorithm: Hybrid Random Walk

- Updating cross-domain links



(a) Route 1 (b) Route 2

$$p_{ij}^{(\mathcal{UP})^+} = \delta \sum_{u_k \in \mathcal{U}} r_{ik}^{(\mathcal{U})} p_{kj}^{(\mathcal{UP})^+} + (1-\delta) \sum_{p_k \in \mathcal{P}} p_{ik}^{(\mathcal{UP})^+} r_{kj}^{(\mathcal{P})}$$

$$p_{ij}^{(\mathcal{UP})^-} = \delta \sum_{u_k \in \mathcal{U}} r_{ik}^{(\mathcal{U})} p_{kj}^{(\mathcal{UP})^-} + (1-\delta) \sum_{p_k \in \mathcal{P}} p_{ik}^{(\mathcal{UP})^-} r_{kj}^{(\mathcal{P})}$$

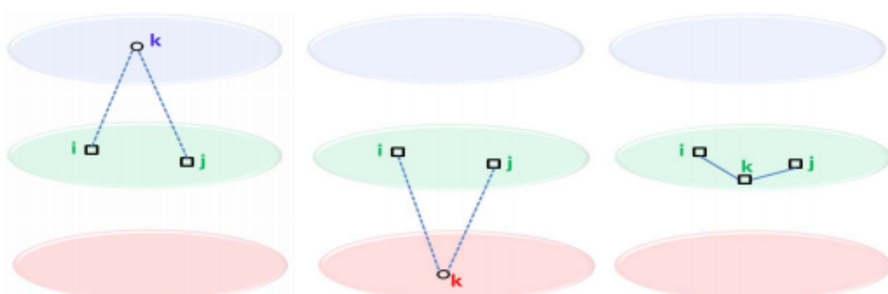
$$p_{ij}^{(\mathcal{UT})^+} = \eta \sum_{u_k \in \mathcal{U}} r_{ik}^{(\mathcal{U})} p_{kj}^{(\mathcal{UT})^+} + (1-\eta) \sum_{t_k \in \mathcal{T}} p_{ik}^{(\mathcal{UT})^+} r_{kj}^{(\mathcal{T})}$$

$$\mathbf{P}^{(\mathcal{UP})^+}(t+1) = \delta \mathbf{R}^{(\mathcal{U})}(t) \mathbf{P}^{(\mathcal{UP})^+}(t) + (1-\delta) \mathbf{P}^{(\mathcal{UP})^+}(t) \mathbf{R}^{(\mathcal{P})}$$

$$\mathbf{P}^{(\mathcal{UP})^-}(t+1) = \delta \mathbf{R}^{(\mathcal{U})}(t) \mathbf{P}^{(\mathcal{UP})^-}(t) + (1-\delta) \mathbf{P}^{(\mathcal{UP})^-}(t) \mathbf{R}^{(\mathcal{P})}$$

$$\mathbf{P}^{(\mathcal{UT})^+}(t+1) = \eta \mathbf{R}^{(\mathcal{U})}(t) \mathbf{P}^{(\mathcal{UT})^+}(t) + (1-\eta) \mathbf{P}^{(\mathcal{UT})^+}(t) \mathbf{R}^{(\mathcal{T})}$$

- Updating within-domain links



(a) Route 1 (b) Route 2 (c) Route 3

$$r_{ij}^{(\mathcal{U})} = \tau^{(\mathcal{P})} \Big(\mu \sum_{p_k \in \mathcal{P}} p_{ik}^{(\mathcal{UP})^+} p_{jk}^{(\mathcal{UP})^+} + (1-\mu) \sum_{p_k \in \mathcal{P}} p_{ik}^{(\mathcal{UP})^-} p_{jk}^{(\mathcal{UP})^-}\Big)$$

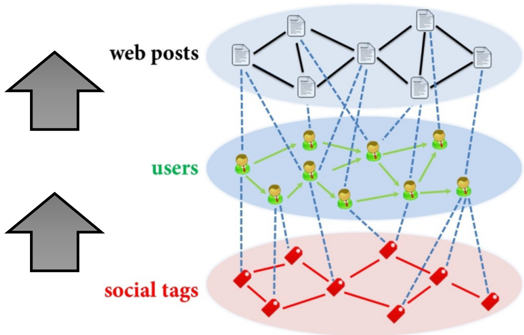
$$+\tau^{(\mathcal{T})} \sum_{t_k \in \mathcal{T}} p_{ik}^{(\mathcal{UT})^+} p_{jk}^{(\mathcal{UT})^+} + \tau^{(\mathcal{U})} \sum_{u_k \in \mathcal{U}} r_{ik}^{(\mathcal{U})} r_{kj}^{(\mathcal{U})} \qquad (12)$$

$$\mathbf{R}^{(\mathcal{U})}(t+1) =$$

$$\tau^{(\mathcal{P})} \Big(\mu \mathbf{P}^{(\mathcal{UP})^+}(t) \mathbf{P}^{(\mathcal{UP})^+}(t)^T + (1-\mu) \mathbf{P}^{(\mathcal{UP})^-}(t) \mathbf{P}^{(\mathcal{UP})^-}(t)^T\Big)$$

$$+\tau^{(\mathcal{T})} \mathbf{P}^{(\mathcal{UT})^+}(t) \mathbf{P}^{(\mathcal{UT})^+}(t)^T + \tau^{(\mathcal{U})} \mathbf{R}^{(\mathcal{U})}(t) \mathbf{R}^{(\mathcal{U})}(t)^T \qquad (13)$$

Results



Comparing with Random Walk with Restarts Models

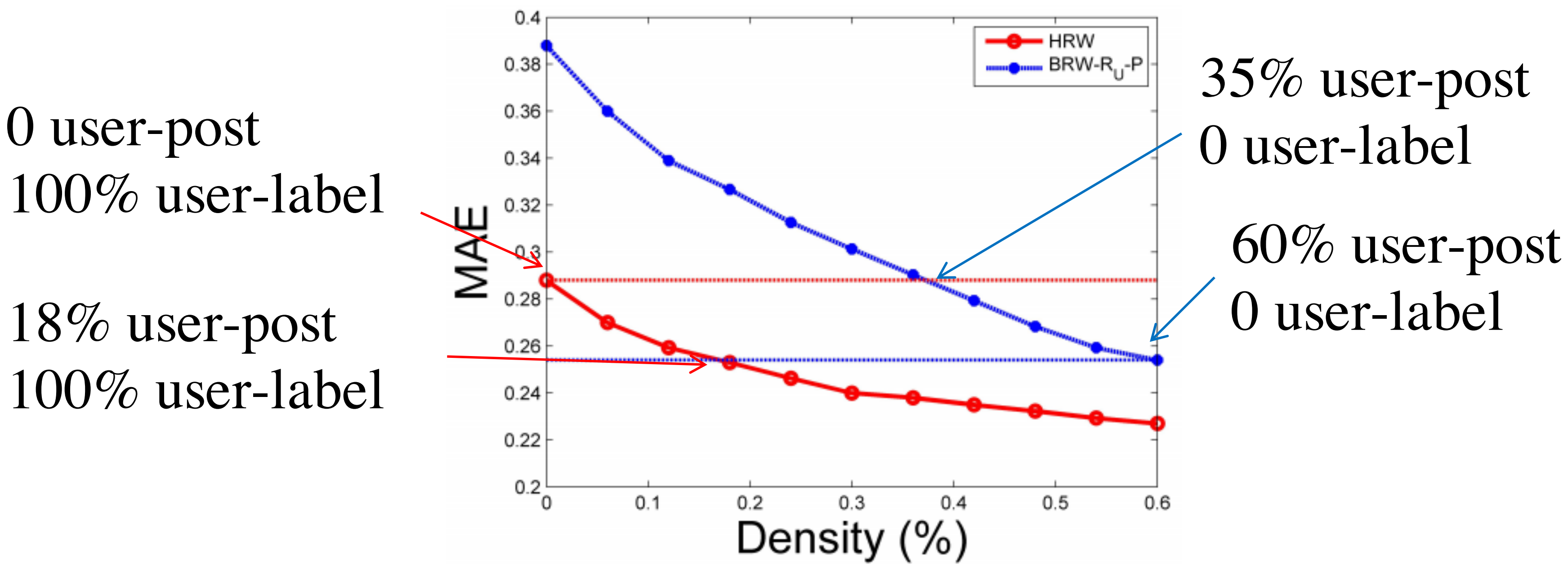
Algorithm	MAE	Precision	Recall	F1	Kendall's $\hat{\tau}$
HRW	**0.227±1.5e-3**	**0.711±1.3e-3**	0.921±1.4e-3	**0.802±1.1e-3**	**0.792±2.5e-3**
BRW-R_U-P (TrustWalker)	0.276±1.1e-3	0.657±7.6e-4	**0.935±9.8e-4**	0.772±7.6e-4	0.774±1.6e-3
BRW-R_U	0.282±5.3e-3	0.655±4.0e-3	0.921±1.2e-2	0.765±7.7e-3	0.725±2.8e-3
BRW-W_U-P	0.292±1.1e-3	0.666±7.0e-4	0.900±5.2e-4	0.765±6.6e-4	0.725±8.5e-4
BRW-W_U (ItemRank)	0.318±1.4e-3	0.671±1.5e-3	0.713±2.4e-3	0.691±1.2e-3	0.661±2.2e-3
BRW-P	0.438±2.6e-4	0.571±3.4e-4	0.499±4.2e-4	0.532±3.2e-4	0.606±2.3e-4

Comparing with Social Recommendation Baselines

Algorithm	MAE	Precision	Recall	F1	Kendall's $\hat{\tau}$
HRW	**0.227±1.5e-3**	**0.711±1.3e-3**	0.921±1.4e-3	**0.802±1.1e-3**	**0.792±2.5e-3**
BRW-R_U-P (TrustWalker) [10]	0.276±1.1e-3	0.657±7.6e-4	0.935±9.8e-4	0.772±7.6e-4	0.774±1.6e-3
BRW-W_U (ItemRank) [8]	0.318±1.4e-3	0.671±1.5e-3	0.713±2.4e-3	0.691±1.2e-3	0.661±2.2e-3
MCF [5]	0.352±2.3e-4	0.592±1.8e-3	**0.951±6.0e-4**	0.730±1.3e-3	0.582±4.3e-4
CF [22]	0.506±3.4e-4	0.552±1.5e-3	0.589±7.2e-4	0.570±1.0e-3	0.540±5.2e-4

Results: Insight

- Knowledge transfer from auxiliary domains improves cold-start users' behavior prediction
  - Using aux. (label) data, saving ***60-70%*** tgt. (post) data



Observation: Multiple Platforms

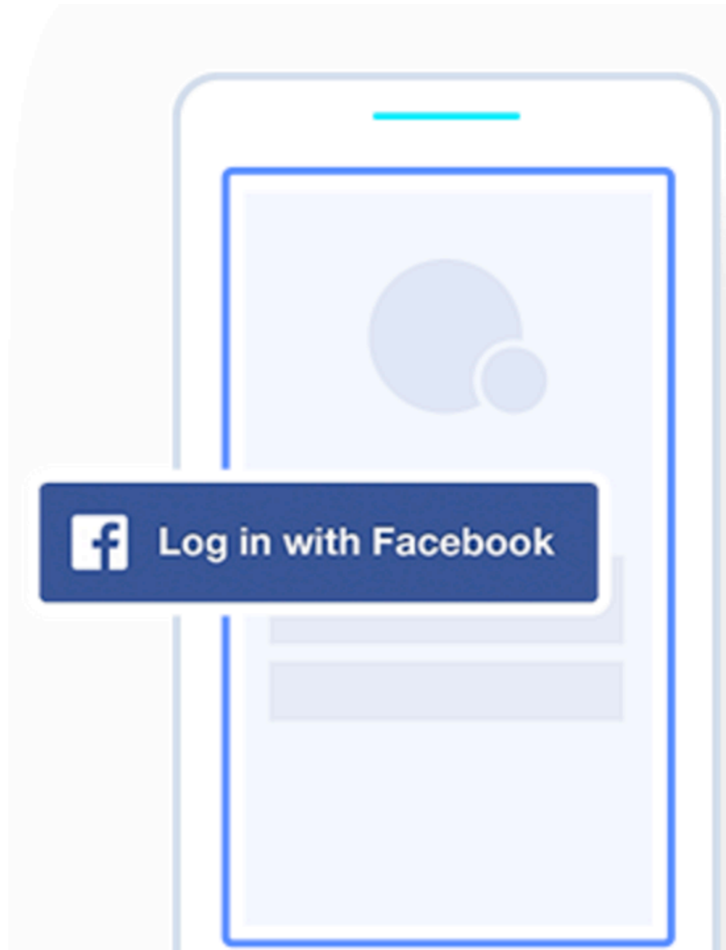
Jiang et al. **Little is Much: Bridging Cross-Platform Behaviors through Overlapped Crowds.** *AAAI*, 2016.

Observation: Cross-Platform

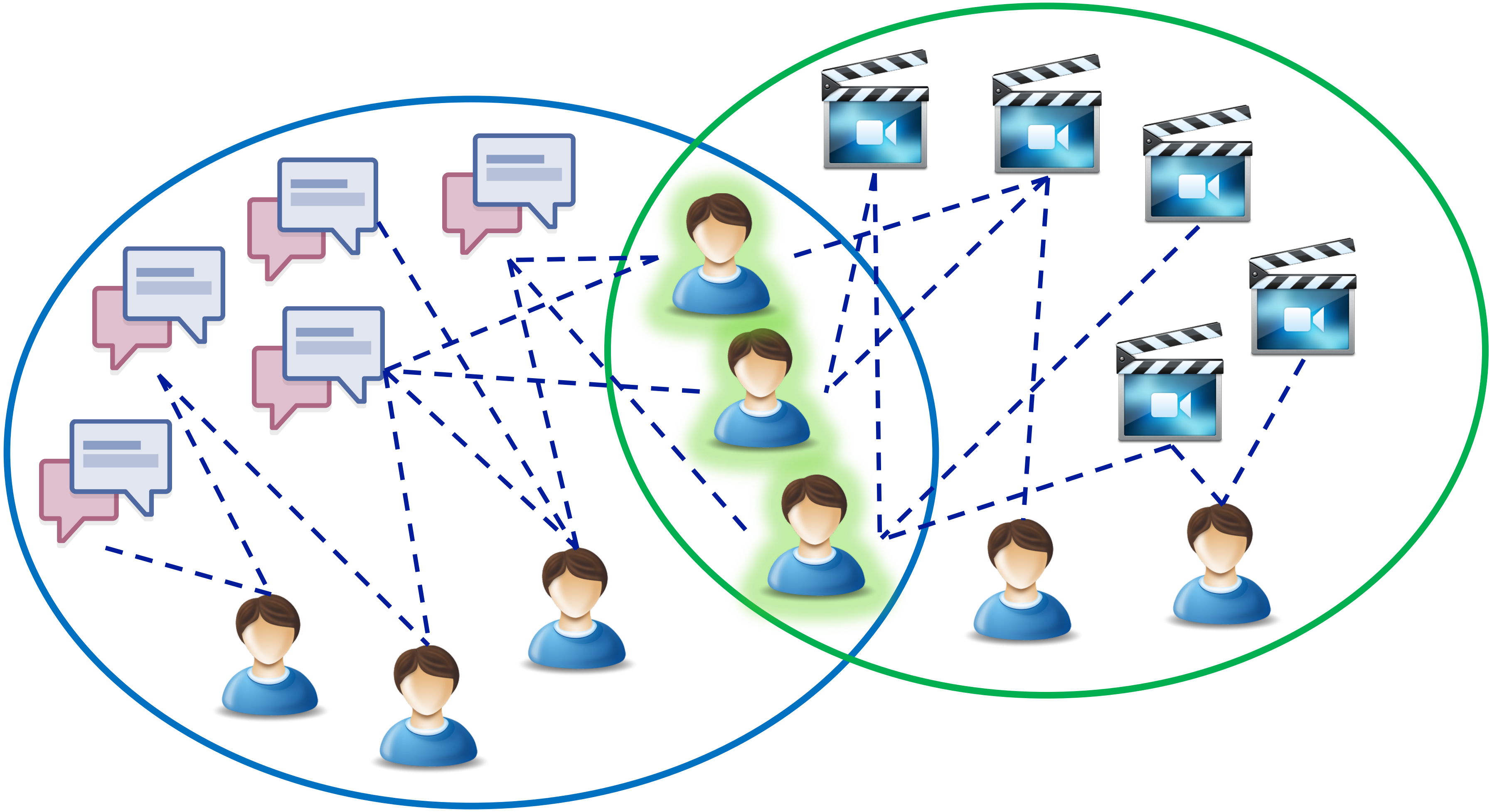
Add Facebook Login to Your App or Website

Facebook Login for Apps is a secure, fast and convenient way for people to log into your app or website.

- iOS
- Android
- Websites or mobile websites
- More platforms



Observation: Partially Overlapped Crowds



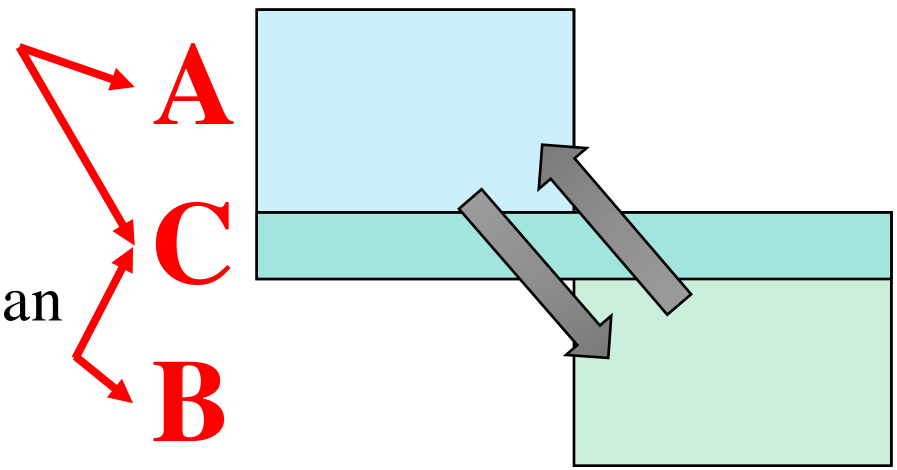
Representation: When NO Transfer

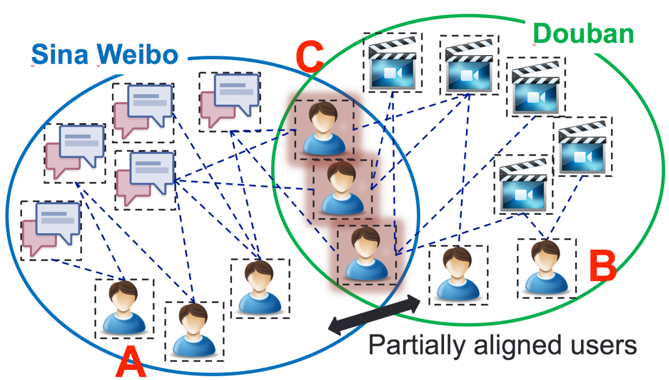
Sina Weibo users

Douban users

Sina Weibo tags/entities

Douban movies/books

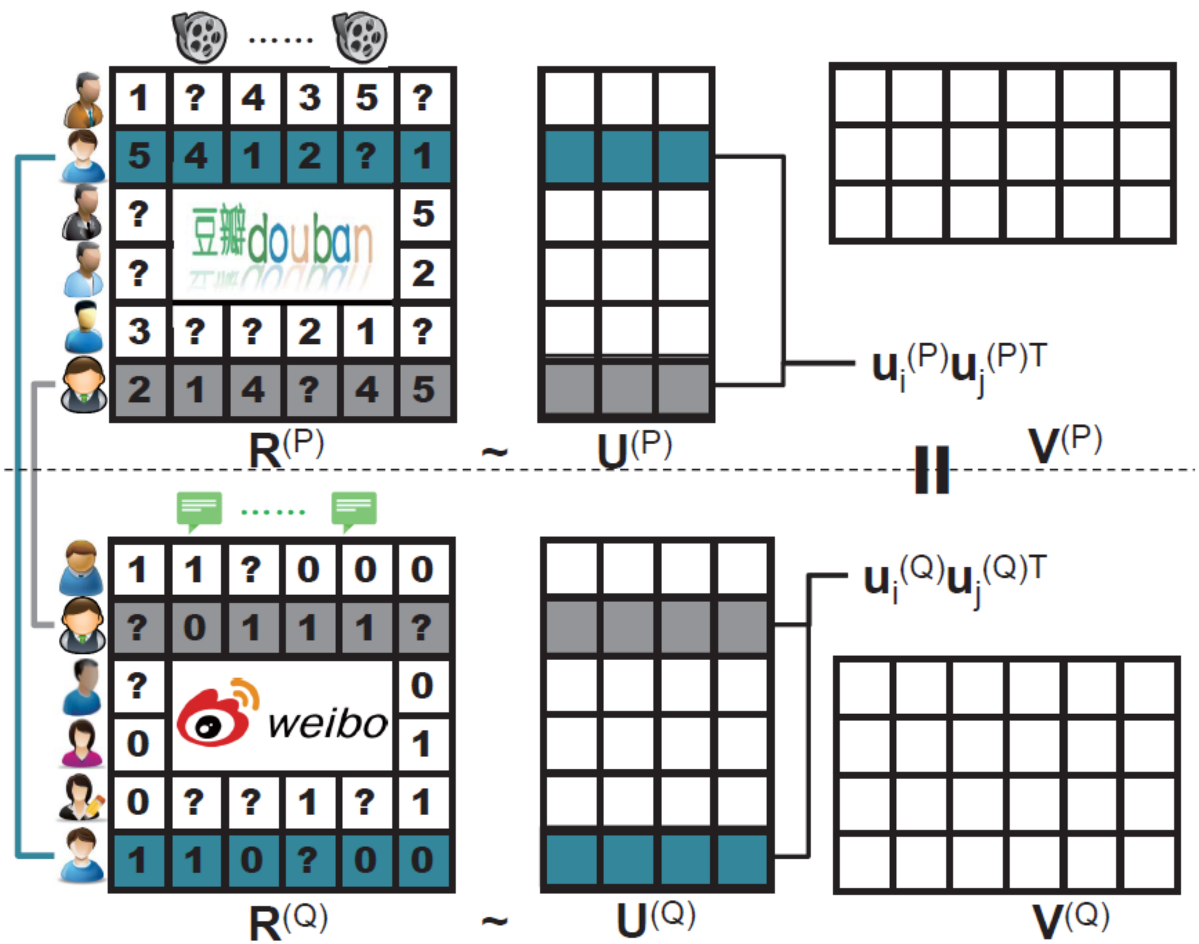




User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A	Auxiliary platform data!	
C	**0.779**	**0.805**
B	1.439	0.640

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	0.429	0.464
C	**0.267**	**0.666**
B	Auxiliary platform data!	

Algorithm: XPTrans



Algorithm: XPTrans

- Input
  - Tgt./Aux. platform P/Q;
  - Behavior data R(P)/R(Q);
  - Observation W(P)/W(Q);
  - Overlapping indicator W(P,Q),
- Output
  - User latent representation U(P)/U(Q);
  - Item latent representation V(P)/V(Q);
  - Missing values in R(P)

- Objective function

Target platform

Auxiliary platform

$$\mathcal{J} = \sum_{i,j} W_{i,j}^{(P)} \left(R_{i,j}^{(P)} - \sum_{r} U_{i,r}^{(P)} V_{r,j}^{(P)} \right)^2 + \lambda \sum_{i,j} W_{i,j}^{(Q)} \left(R_{i,j}^{(Q)} - \sum_{r} U_{i,r}^{(Q)} V_{r,j}^{(Q)} \right)^2 + \mu \sum_{i_1,j_1,i_2,j_2} W_{i_1,j_1}^{(P,Q)} W_{i_2,j_2}^{(P,Q)} \left(A_{i_1,i_2}^{(P)} - A_{j_1,j_2}^{(Q)} \right)^2$$

Overlapping user similarity
(Pair-wise regularization)

Results: Leveraging Auxiliary Platform Data

NO Transfer

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A	Auxiliary platform data!	
C	0.779	0.805
B	**1.439**	**0.640**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.429**	**0.464**
C	0.267	0.666
B	Auxiliary platform data!	

Transfer via the Same Latent Space

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A		
C	0.757	0.811
B	**1.164 (-19%)**	**0.702 (+9.7%)**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.411 (-4.2%)**	**0.487 (+5.0%)**
C	0.256	0.681
B		

Results: Leveraging Different Latent Spaces

Transfer via the Same Latent Space

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A		
C	0.757	0.811
B	**1.164**	**0.702**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.411**	**0.487**
C	0.256	0.681
B		

Transfer via Different Latent Spaces

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A		
C	0.715	0.821
B	**0.722 (-38%)**	**0.820 (+17%)**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.374 (-11%)**	**0.533 (+12%)**
C	0.236	0.705
B		

Results: Where Amazing Happens

NO Transfer

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A	Auxiliary platform data!	
C	0.779	0.805
B	**1.439**	**0.640**

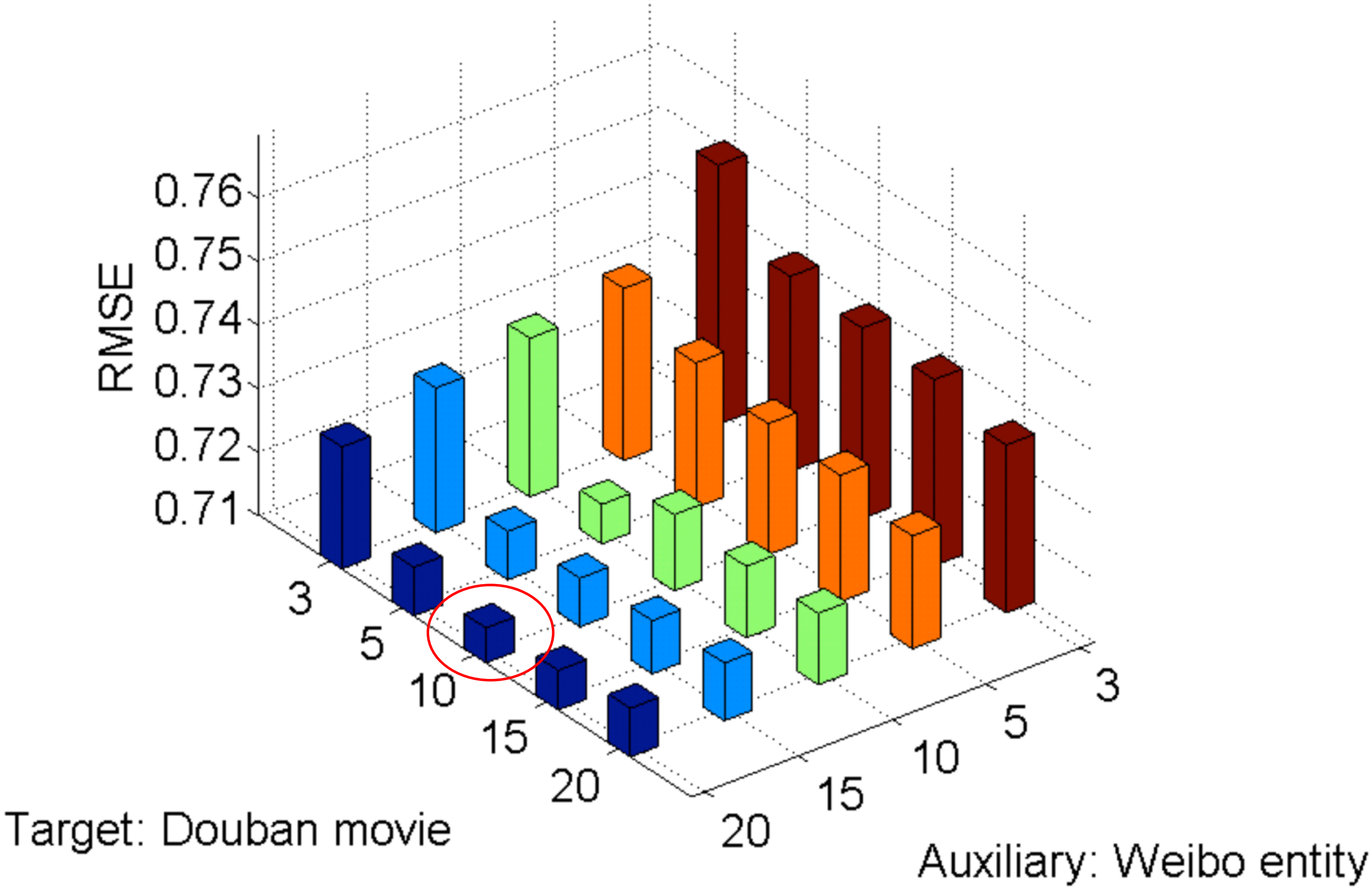
User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.429**	**0.464**
C	0.267	0.666
B	Auxiliary platform data!	

Transfer via Different Latent Spaces

User set	Weibo tweet entity to Douban movie	
	RMSE	MAP
A		
C	0.715	0.821
B	**0.722**	**0.820**

User set	Douban book to Weibo social tag	
	RMSE	MAP
A	**0.374**	**0.533**
C	0.236	0.705
B		

Results: Different Sizes of Latent Spaces



Summary

- Like, Reply, Share, Retweet, Favorite, Comment …
- Memory based social recommenders
  - TidalTrust, MoleTrust, TrustWalker
- Model based social recommenders
  - SoRec, "Social Trust" Ensemble, SoReg
- **Observations, Representations, Models**
  - **ContextMF:** Social contexts (preference & influence)
  - **FEMA:** Spatiotemporal contexts (multidimensional)
  - **HybridRW:** Cross-domain behavior modeling
  - **XPTrans:** Cross-platform behavior modeling

Acknowledgement

NSF
NSFC
National Natural Science Foundation of China
Carnegie Mellon University
UNIVERSITY OF MARYLAND 1856
Microsoft Research
微软亚洲研究院
Tencent 腾讯

References

D. Blei, A. Ng, and M. Jordan. "Latent dirichlet allocation." JMLR, 2003.

J. Herlocker, J. Konstan, L. Terveen, J. Riedl. "Evaluating collaborative filtering recommender systems." ACM TOIS, 2004.

Y. Koren, R. Bell, C. Volinsky. "Matrix factorization techniques for recommender systems." Computer, 2009.

Y. Koren. "Factorization meets the neighborhood: A multifaceted collaborative filtering model." KDD, 2008.

Y. Koren. "Collaborative filtering with temporal dynamics." CACM, 2010.

M. Balabanovic and Y. Shoham. "FAB: Content-based, collaborative recommendation." CACM, 1997.

N. Liu and Q. Yang. "Eigenrank: A ranking-oriented approach to collaborative filtering." SIGIR, 2008.

N. Liu, M. Zhao, and Q. Yang. "Probabilistic latent preference analysis for collaborative filtering." CIKM, 2009.

References

H. Ma, H. Yang, M. Lyu, and I. King. "Sorec: Social recommendation using probabilistic matrix factorization." CIKM, 2008.

H. Ma, T. Zhou, M. Lyu, and I. King. "Improving recommender systems by incorporating social contextual information." ACM TOIS, 2011.

H. Ma, D. Zhou, C. Liu, M. Lyu, and I. King. "Recommender systems with social regularization." WSDM, 2011.

J. Leskovec, A. Singh, and J. Kleinberg. "Patterns of influence in a recommendation network." PAKDD, 2006.

P. Massa and A. Paolo. "Trust-aware recommender systems." RecSys, 2007.

M. Jamali and E. Martin. "TrustWalker: A random walk model for combining trust-based and item-based recommendation." KDD, 2009.

H. Ma, I. King, and M. Lyu. "Learning to recommend with social trust ensemble." SIGIR, 2009.

H. Ma. I. King, and M. Lyu. "Learning to recommend with explicit and implicit social relations." ACM TIST, 2011.

References

M. Faloutsos, P. Faloutsos, and C. Faloutsos. "On power-law relationships of the internet topology." SIGCOMM, 1999.

A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Weiner. "Graph structure in the web." Computer Networks, 2000.

F. Chung and L. Lu. "The average distances in random graphs with given expected degrees." PNAS, 2002.

J. Kleinberg. "Authoritative sources in a hyperlinked environment." JACM, 1999.

H. Kwak, C. Lee, H. Park, and S. Moon. "What is Twitter, a social network or a news media?" WWW, 2010.

B. Hooi, H.A. Song, A. Beutel, N. Shah, K. Shin, and C. Faloutsos. "Fraudar: Bounding graph fraud in the face of camouflage." KDD, 2016.

C. Aggarwal and J. Han. "Frequent pattern mining." Springer, 2014.

J. Han, J. Pei, B. Mortazavi-Asl, Q. Chen, U. Dayal, and M. Hsu. "FreeSpan: Frequent pattern-projected sequential pattern mining." KDD, 2000.

References

X. Yan and J. Han. "gspan: Graph-based substructure pattern mining." ICDM, 2003.

X. Yan and J. Han. "CloseGraph: Mining closed frequent graph patterns." KDD, 2003.

Y. Sun, J. Han, X. Yan, P.S. Yu, and T. Wu. "PathSim: Meta path-based top-k similarity search in heterogeneous information networks." VLDB, 2011.

Y. Sun, Y. Yu, and J. Han. "Ranking-based clustering of heterogeneous information networks with star network schema." KDD, 2009.

Y. Sun, J. Han, P. Zhao, Z. Yin, H. Cheng, and T. Wu. "RankClus: Integrating clustering with ranking for heterogeneous information network analysis." EDBT, 2009.

Y. Sun, R. Barber, M. Gupta, C. Aggarwar, and J. Han. "Co-author relationship prediction in heterogeneous bibliographic networks." ASONAM, 2011.

A. El-Kishky, Y. Song, C. Wang, C.R. Voss, and J. Han. "Scalable topical phrase mining from text corpora." VLDB, 2014.

J. Liu, J. Shang, C. Wang, X. Ren, and J. Han. "Mining quality phrases from massive text corpora." SIGMOD, 2015.

References

X. Ren, A. El-Kishky, C. Wang, F. Tao, C.R. Voss, and J. Han. "Effective entity recognition and typing by relation phrase-based clustering." KDD, 2015.

X. Ren, W. He, M. Qu, C.R. Voss, H. Ji, and J. Han. "Label noise reduction in entity typing by heterogeneous partial-label embedding." KDD, 2016.

C. Wang, M. Danilevsky, N. Desai, Y. Zhang, P. Nguyen, T. Taula, and J. Han. "A phrase mining framework for recursive construction of a topical hierarchy." KDD, 2013.

E.E. Papalexakis, C. Faloutsos, N.D. Sidiropoulos. "ParCube: Sparse parallelizable tensor decompositions." PKDD, 2012.

D. Koutra, U. Kang, J. Vreeken, and C. Faloutsos. "VOG: Summarizing and understanding large graphs." SDM, 2014.

R. Gupta, A. Halevy, X. Wang, S.E. Whang, and F. Wu. "Biperpedia: An ontology for search applications." VLDB, 2014.

M. Yahya, S. Whang, R. Gupta, and A. Halevy. "ReNoun: Fact extraction for nominal attributes." EMNLP, 2014.

A. Halevy, N. Noy, S. Sarawagi, S.E. Whang, and X. Yu. "Discovering structure in the universe of attribute names." WWW, 2016.

References

Q. Li, Y. Li, J. Gao, B. Zhao, W. Fan, and J. Han. "Resolving conflicts in heterogeneous data by truth discovery and source reliability estimation." SIGMOD, 2014.

Q. Li, Y. Li, J. Gao, L. Su, B. Zhao, M. Demirbas, W. Fan, and J. Han. "A confidence-aware approach for truth discovery on long-tail data." VLDB, 2014.

F. Ma, Y. Li, Q. Li, M. Qiu, J. Gao, S. Zhi, L. Su, B. Zhao, H. Ji, and J. Han. "Faitcrowd: Fine grained truth discovery for crowdsourced data aggregation." KDD, 2015.

Y. Li, J. Gao, C. Meng, Q. Li, L. Su, B. Zhao, W. Fan, and J. Han. "A survey on truth discovery." KDD Explorations Newsletter, 2016.

S. Zhi, B. Zhao, W. Tong, J. Gao, D. Yu, H. Ji, and J. Han. "Modeling truth existence in truth discovery." KDD, 2015.

S. Kumar, R. West, and J. Leskovec. "Disinformation on the Web: Impact, characteristics, and detection of Wikipedia hoaxes." WWW, 2016.

S. Kumar, F. Spezzano, and V.S. Subrahmanian. "Identifying malicious actors on social media." ASONAM, 2016. (tutorial)

Thank you!

Data-Driven Behavioral Analytics:
Observations, Representations and Models