SOCIAL RECOMMENDATION ACROSS MULTIPLE RELATIONAL DOMAINS

Meng Jiang

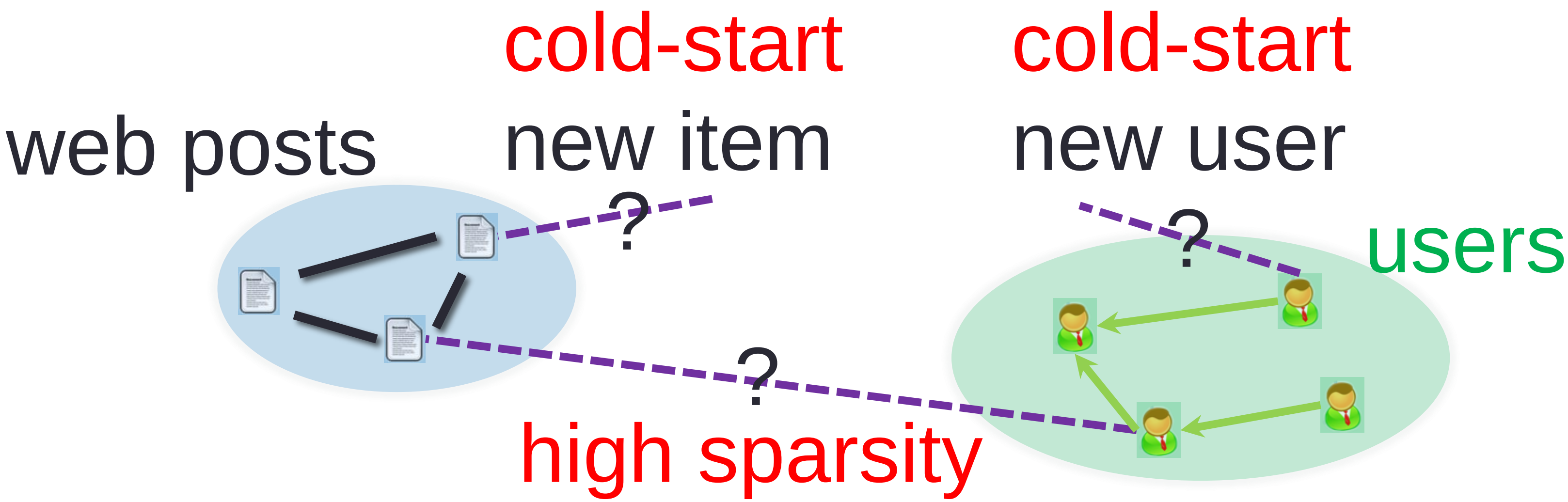
Joint work with Peng Cui, Fei Wang, Qiang Yang,
Wenwu Zhu and Shiqiang Yang
November 1, 2012 – Maui, HI, USA

Recommender Systems

Predict missing “user-item” links

Challenge:

Cold-start and extremely high sparsity



OUTLINE

1. Background

2. The Framework

3. HRW Algorithm

4. Experiments

5. Insights

Multiple Domains

- User label domain

Choose < 10 from 200+ labels like ‘iPhone fan’

Peng Cui
Haidian, Beijing
Company: Tsinghua

User labels (5)
Tsinghua, Ph.D., World Wide Web, Social Network, Social Media

Meng Jiang
Haidian, Beijing
University: Tsinghua

User labels (9)
Chinese food, World Wide Web, Social Network, Data Mining, Liverpool Football Club, NBA, Humors, Sports, Ph.D. Candidates

Multiple Domains

- Interest group domain

Interest Groups (2)

Tsinghua University

I love sing!

Interest Groups (3)

Tsinghua University

Social Media & Reputation

World Wide Web Team

Our Goals

- Given: Links on social networks
- Find: A framework that use auxiliary knowledge in multiple domains to best predict “user-item” (target) links when the training set is too small.

- Goals:
  - G1. Understand link formations on social networks
  - G2. A social network framework with multiple domains
  - G3. Solve the cold-start problem

Challenges: Multiple Domains

- Relational
  - Within-domain links and cross-domain links
- Heterogeneous
  - Different types of item domains
- Sparse
  - Different sparsity levels

OUTLINE

1. Background

2. The Framework

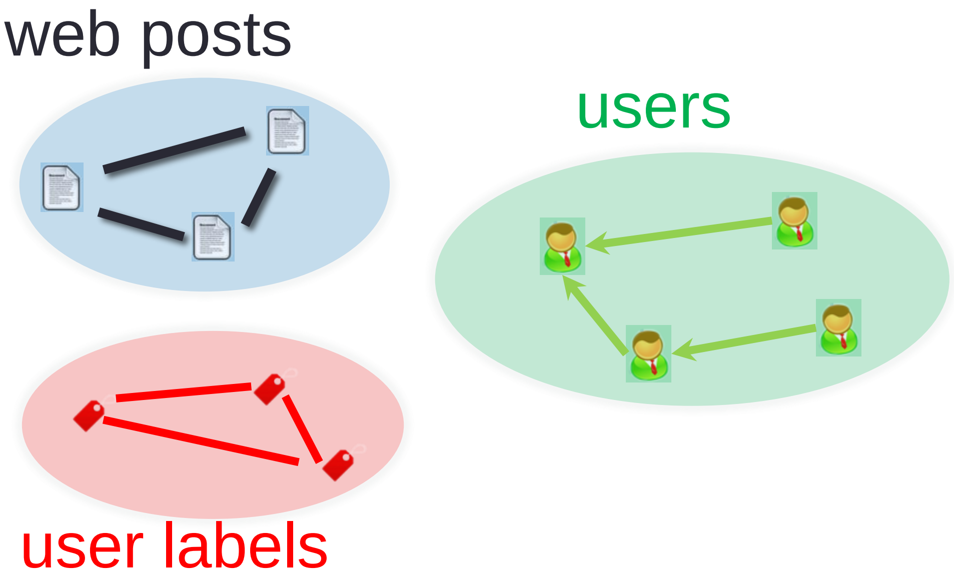
3. HRW Algorithm

4. Experiments

5. Insights

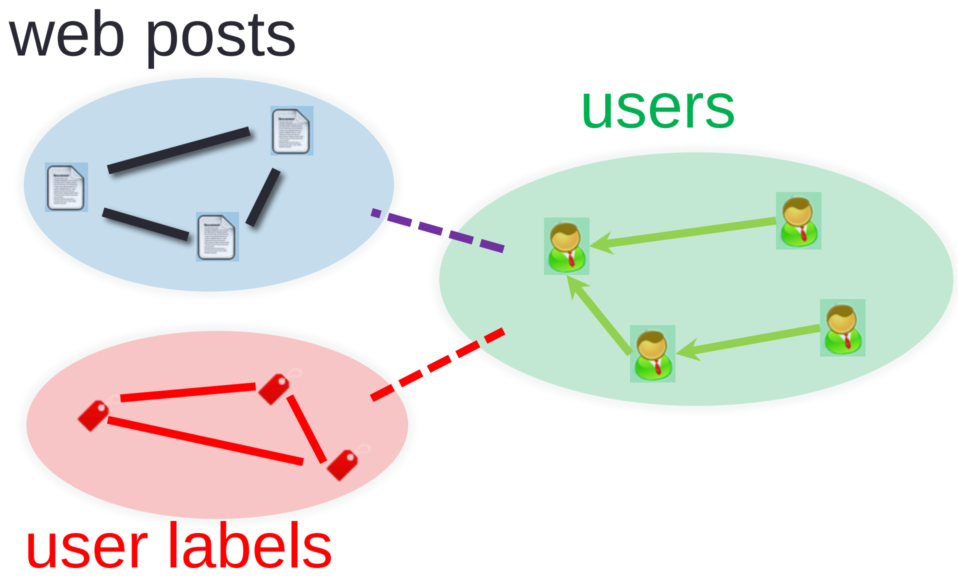
Reframe Social Networks

- We have user-user, post-post and label-label links (social relation + item similarity).



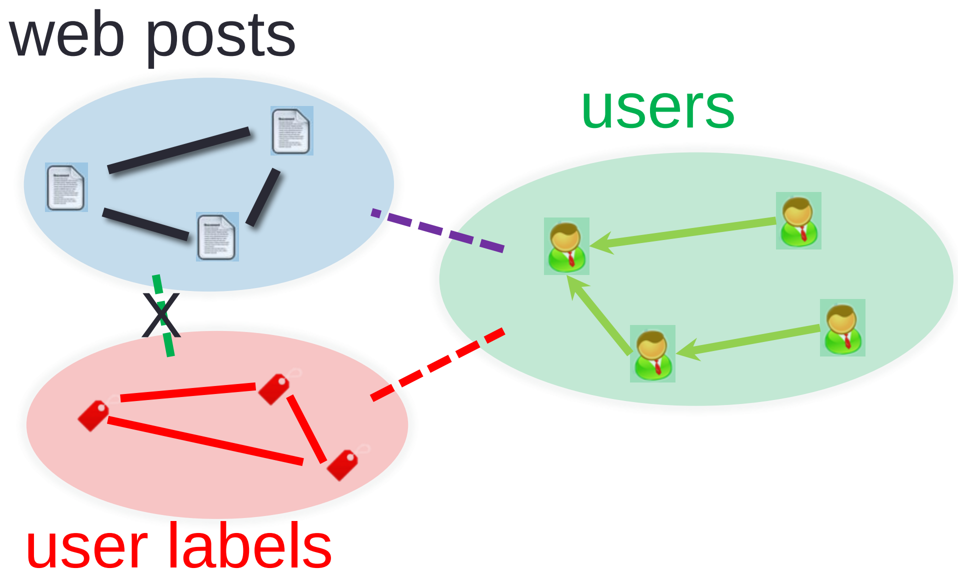
Reframe Social Networks

- We have user-post and user-label links.



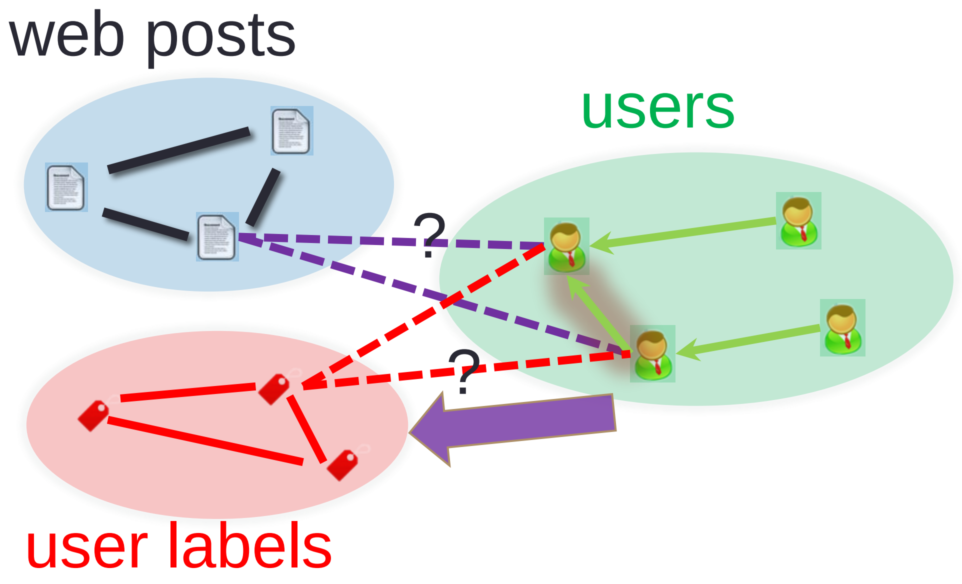
Reframe Social Networks

- No relations between item domains.
- No post-label links in nature.



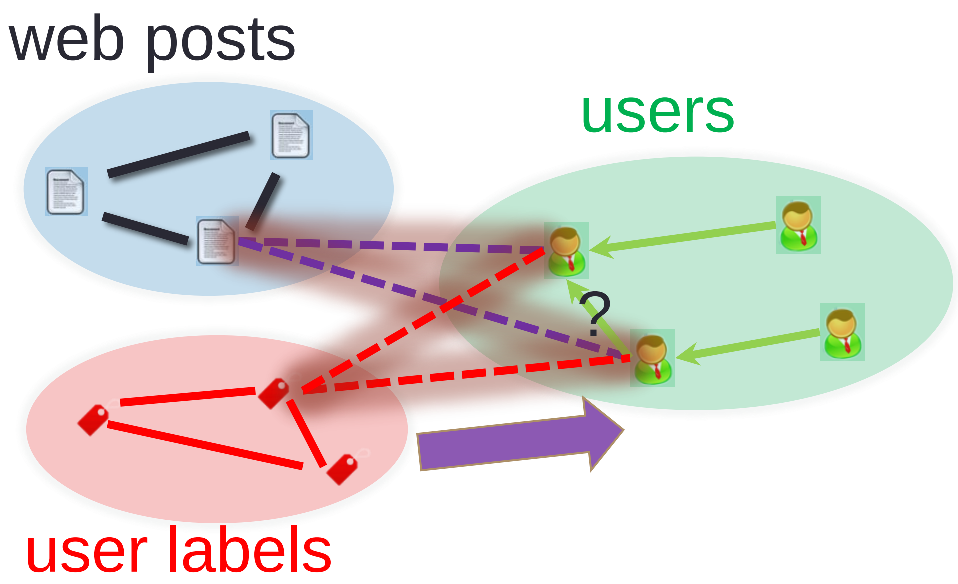
Reframe Social Networks

- Stronger social relations help collaborate user-item links.



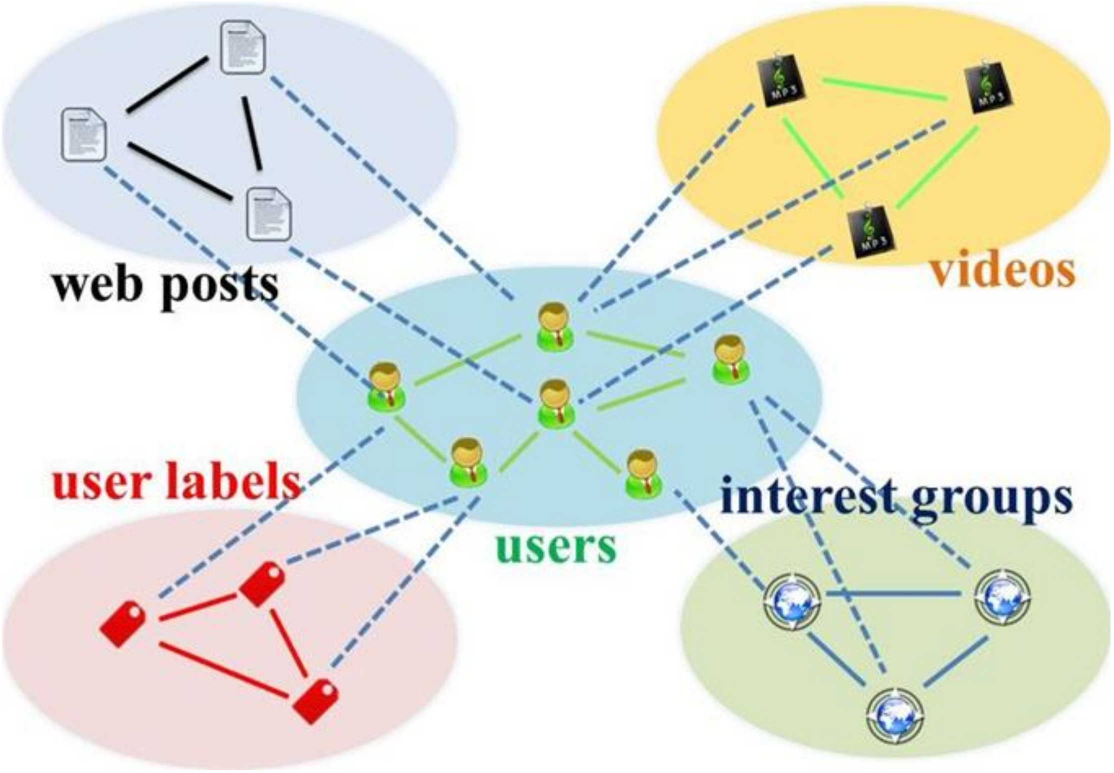
Reframe Social Networks

- More collaborating in user-item links strengthen the social relations.



Star-structured Graph

- Key idea: use “social relation” domain as bridge

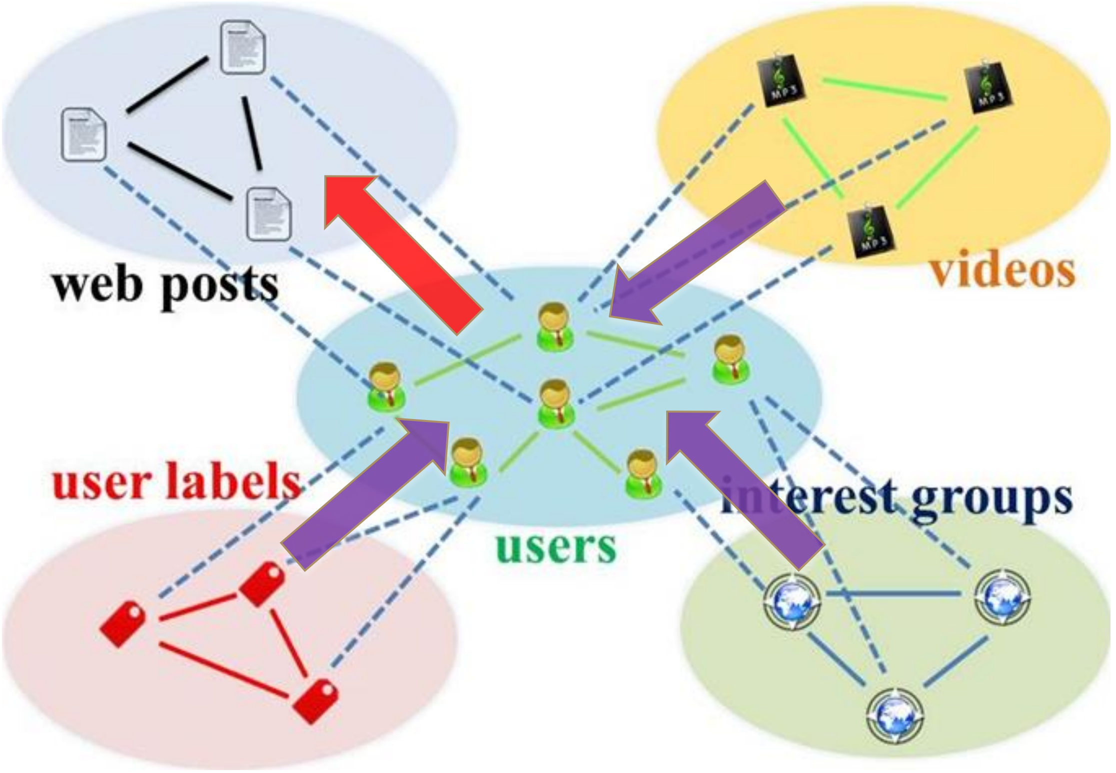


OUTLINE

1. Background
2. The Framework
3. **HRW Algorithm**
4. Experiments
5. Insights

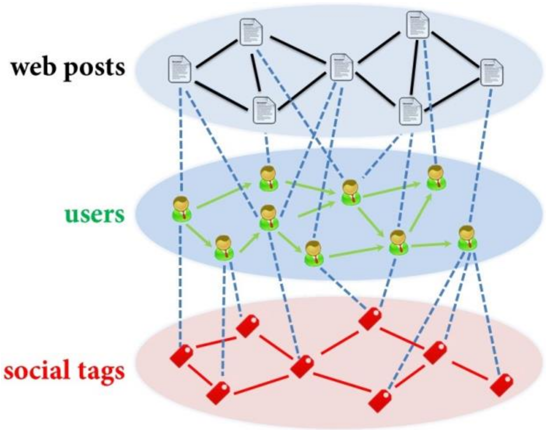
Star-structured Graph

- Method: Transfer learning + Random walk with restarts

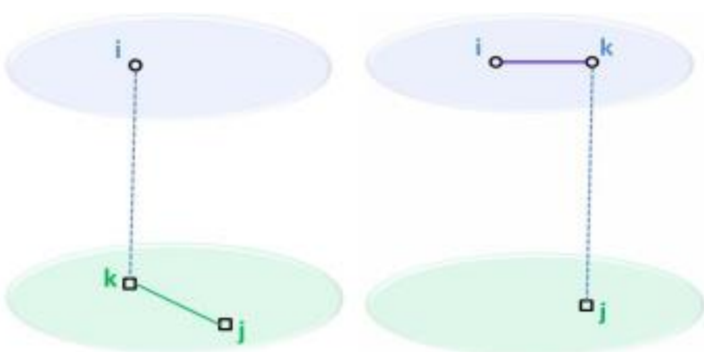


Hybrid Random Walk

- On second-order star-structured graph



- Update cross-domain links



(a) Route 1 (b) Route 2

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

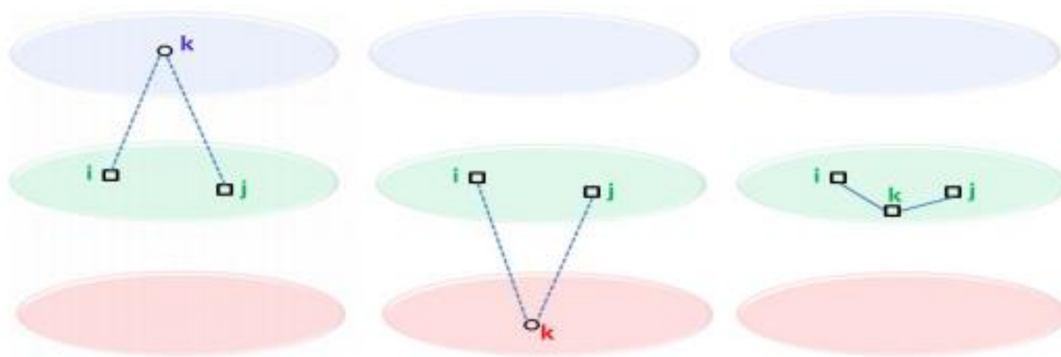
$$p_{ij}^{(\mathcal{UP})^-} = \delta \sum_{u_k \in \mathcal{U}} r_{ik}^{(\mathcal{U})} p_{kj}^{(\mathcal{UP})^-} + (1-\delta) \sum_{p_k \in \mathcal{P}} p_{ik}^{(\mathcal{UP})^-} r_{kj}^{(\mathcal{P})}$$

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

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

Hybrid Random Walk

- Update within-domain links



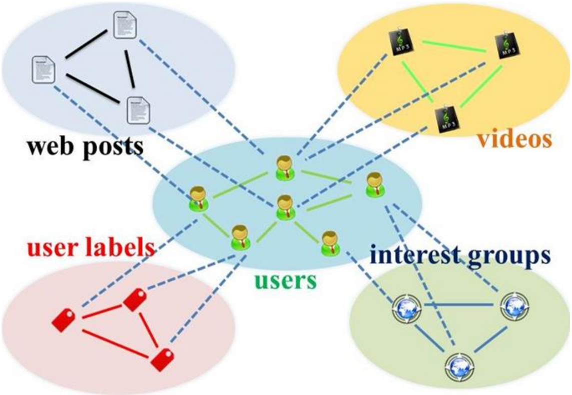
(a) Route 1 (b) Route 2 (c) Route 3

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

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

Hybrid Random Walk

- On high-order star-structured graph



$$
\begin{aligned}
\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t+1) &= \delta_i \mathbf{R}^{(\mathcal{U})}(t)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t) + (1-\delta_i)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t)\mathbf{R}^{(\mathcal{D}_i)} \\
\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t+1) &= \delta_i \mathbf{R}^{(\mathcal{U})}(t)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t) + (1-\delta_i)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t)\mathbf{R}^{(\mathcal{D}_i)} \\
\mathbf{R}^{(\mathcal{U})}(t+1) &= \sum_{\mathcal{D}_i \in \mathcal{D}} \tau_i \mu_i \mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t)^T \\
&\quad + \sum_{\mathcal{D}_i \in \mathcal{D}} \tau_i (1-\mu_i)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t)^T \\
&\quad + \tau^{(\mathcal{U})}\mathbf{R}^{(\mathcal{U})}(t)\mathbf{R}^{(\mathcal{U})}(t)^T
\end{aligned}
$$

OUTLINE

Data Set

- Tencent Weibo (January 2011)

Domain	Size	Cross-domain links	
		Accept	Refuse
User	53.4K	—	—
Web post	142K	1.47M (0.02%)	3.40M (0.04%)
User label	111	330K (5.57%)	—

Good to Transfer?

- Comparative Algorithms (RWR)
  - $W^{(P)}$: Use web post similarity?
  - $W^{(U)}$: Use social relation?
  - $R^{(U)}$: Update tie strength?
  - $W^{(T)}$: Use user label similarity?

Algorithm	$\mathbf{R}^{(\mathcal{U})}$	$\mathbf{W}^{(\mathcal{U})}$	$\mathbf{W}^{(\mathcal{P})}$	$\mathbf{W}^{(\mathcal{T})}$
HRW	√	√	√	√
BRW-R_U-P (TrustWalker)	√	√	√	×
BRW-R_U	√	√	×	×
BRW-W_U-P	×	√	√	×
BRW-W_U (ItemRank)	×	√	×	×
BRW-P	×	×	√	×

Good to Transfer!

- Compare with RWR 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

- Compare with Baselines

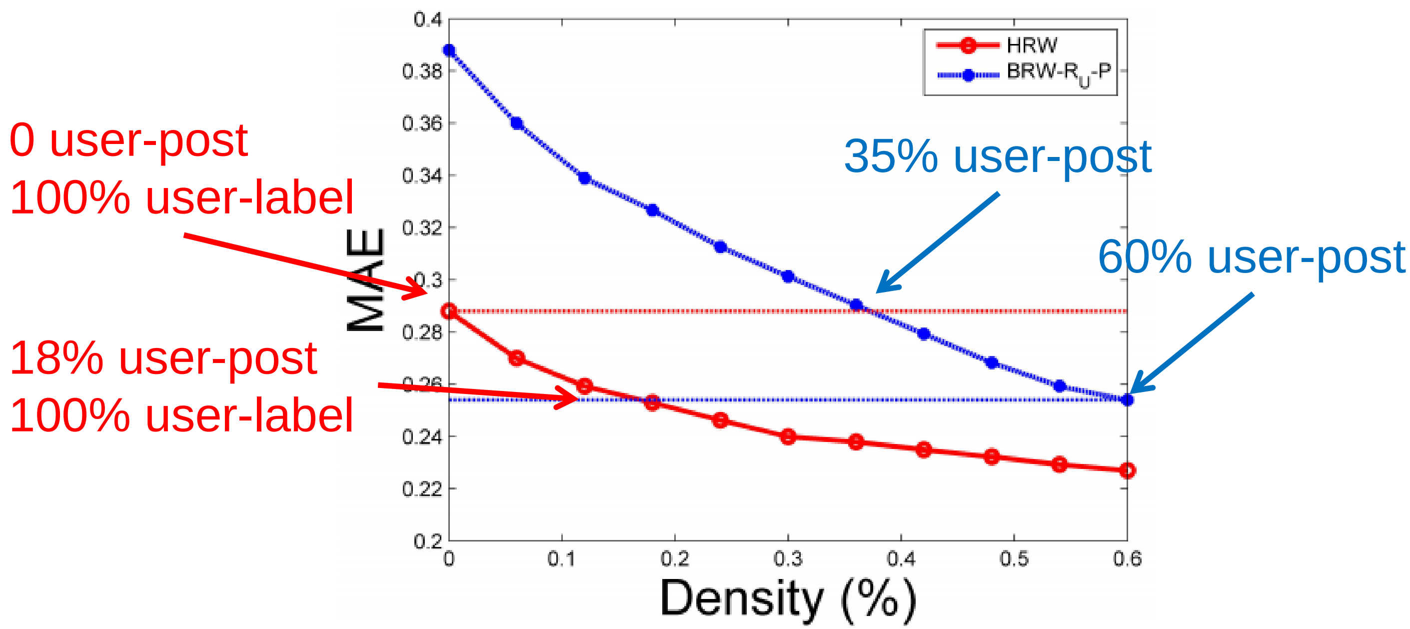
Algorithm	MAE	Precision	Recall	F1	Kendall's $\hat{\tau}$
HRW	**0.227±1.5e-3**	**0.711±1.3e-3**	0.921±1.4e-3	**0.802±1.1e-3**	**0.792±2.5e-3**
BRW-R_U-P (TrustWalker) [10]	0.276±1.1e-3	0.657±7.6e-4	0.935±9.8e-4	0.772±7.6e-4	0.774±1.6e-3
BRW-W_U (ItemRank) [8]	0.318±1.4e-3	0.671±1.5e-3	0.713±2.4e-3	0.691±1.2e-3	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

OUTLINE

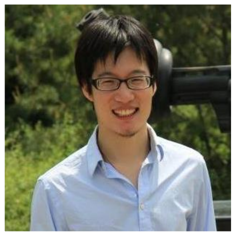
1. Background
2. The Framework
3. HRW Algorithm
4. Experiments
5. **Insights**

Insights

- If we do transfer (from user-label domain), we need only ~30% to reach the same performance.
- Advice: build more apps for new users to give more info.



Questions?





Meng Jiang
mjiang89@gmail.com
http://www.meng-jiang.com