BEHAVIORAL MODELING IN SOCIAL NETWORKS
FROM MICRO TO MACRO

Meng Jiang, *University of Illinois at Urbana-Champaign*
Peng Cui, *Tsinghua University*
http://www.meng-jiang.com/tutorial-icdm15.html

ICDM 2015 TUTORIAL
Atlantic City, NJ

清華大學
Tsinghua University

ILLINOIS
1867
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Thanks to

NSFC
National Natural Science Foundation of China

mda
Media Development Authority Singapore

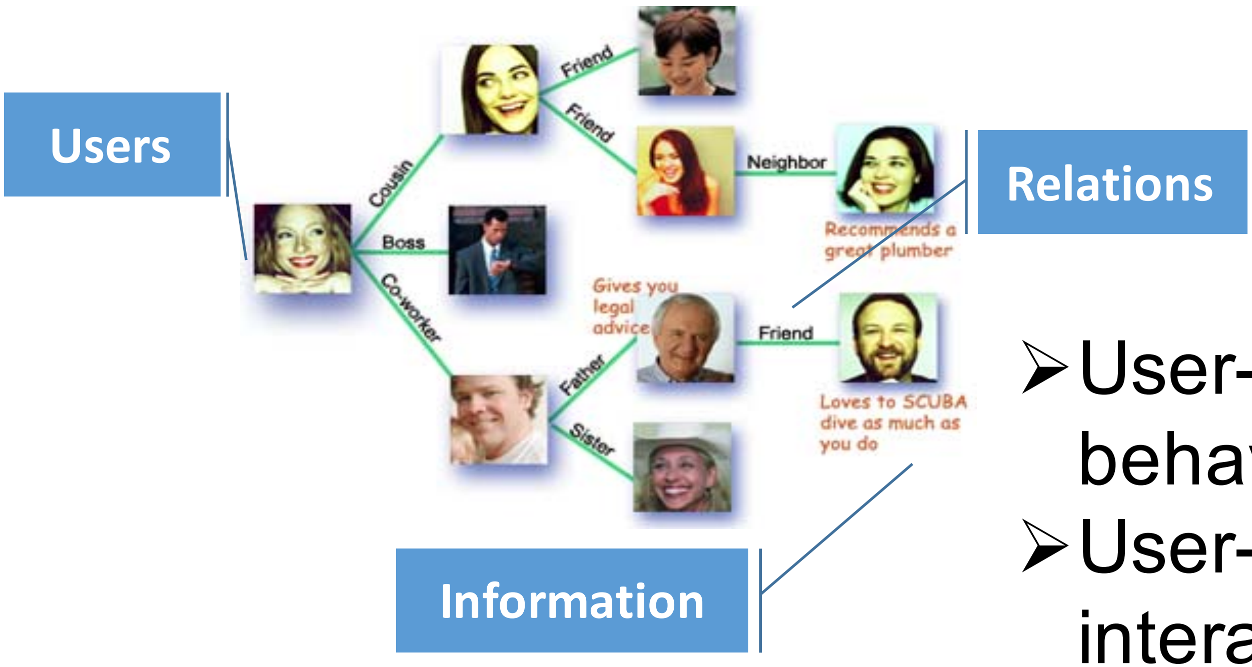
TSINGHUA UNIVERSITY- TENCENT JOINT LABORATORY

National Natural Science Foundation of China,
No. 61370022, No. 61210008, No. 61303075
International Science & Technology Cooperation Program of China,
No. 2013DFG12870
National Program on Key Basic Research Project,
No. 2011CB302206, No. 2015CB352300
NExT Research Center funded by MDA, Singapore,
WBS:R-252-300-001-490
Tsinghua-Tencent Joint Laboratory

INTRODUCTION

Behavior Modeling in Social Networks

User Behavior



- User-user interaction behaviors
- User-information interaction behaviors

User behavior is a fundamental element in social networks

Behavioral Modeling

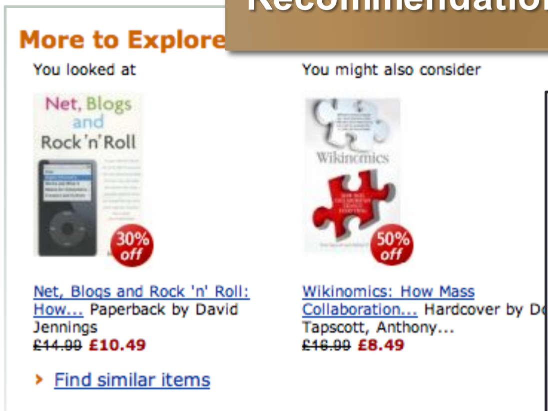
Understanding

Predicting

Intervening

Applications of Behavioral Modeling

Recommendation



amazon Google
twitter YouTube

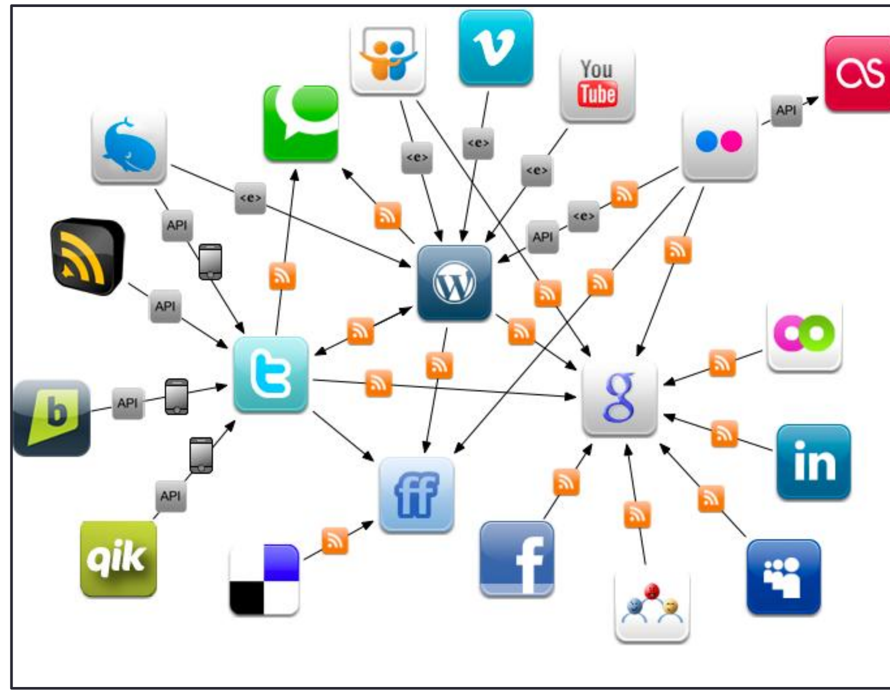
Targeting AD

Anti Spam

Scientific Significance of Behavioral Modeling

Physical World

Online Social Networks



The human behaviors are broadly and deeply recorded in an unprecedented level.

This is the first time that we can get insights of human behaviors and the society from large scale real data.

Six Disruptive Basic Research Areas, by DOD

Basic Research Areas

- Six Disruptive Basic Research Areas
  - Engineered Materials (metamaterials and plasmonics)
  - Quantum Information and Control
  - Cognitive Neuroscience
  - Nanoscience and Nanoengineering
  - Synthetic Biology
  - Computational Modeling of Human and Social Behavior

VI. Computational Models of Human Behavior

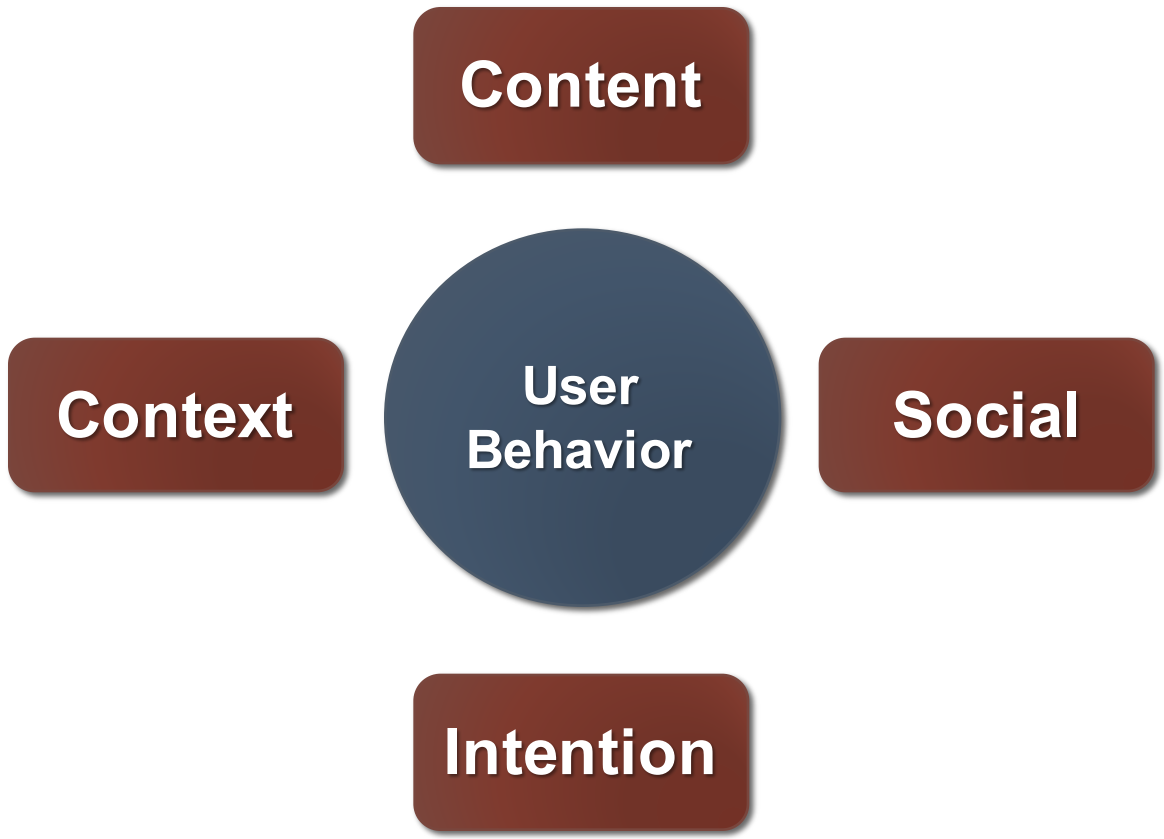
A fundamental understanding and predictive capability of human behavior dynamics from individuals to societies.

- ***Enabled capabilities***
  - *Predictive models supporting strategic, operational, and tactical decision making and planning*
  - *Real time cultural situational awareness*
  - *Immersive training and mission rehearsal*
  - *Cross cultural coalition building*

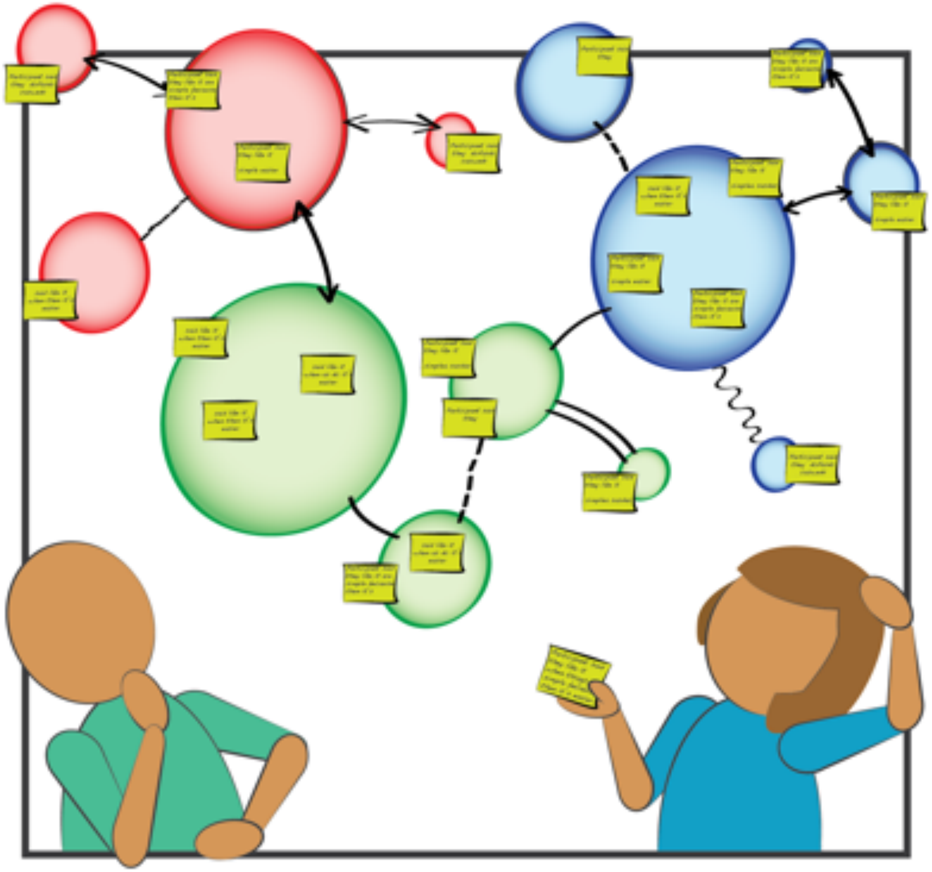
- ***Key research challenges:***
  - *Conflicting theories*
  - *Data management and fusion*
  - *Mathematical complexity*
  - *Validation of models*

- ***Measures of success***
  - *Early success of simple models*
  - *Success of social network analysis*
  - *Prediction of crowd tipping points*

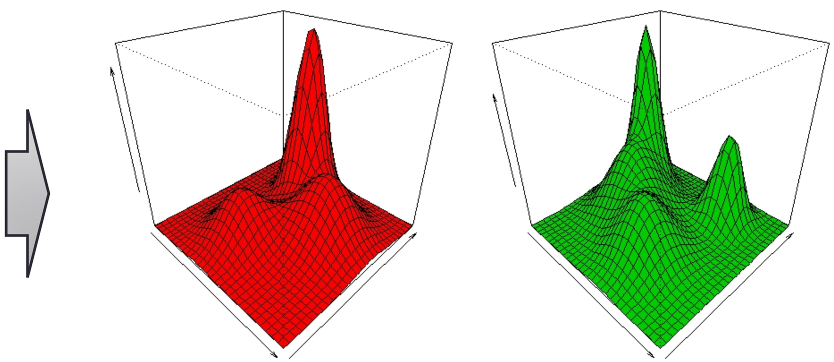
User Behaviors are Complex



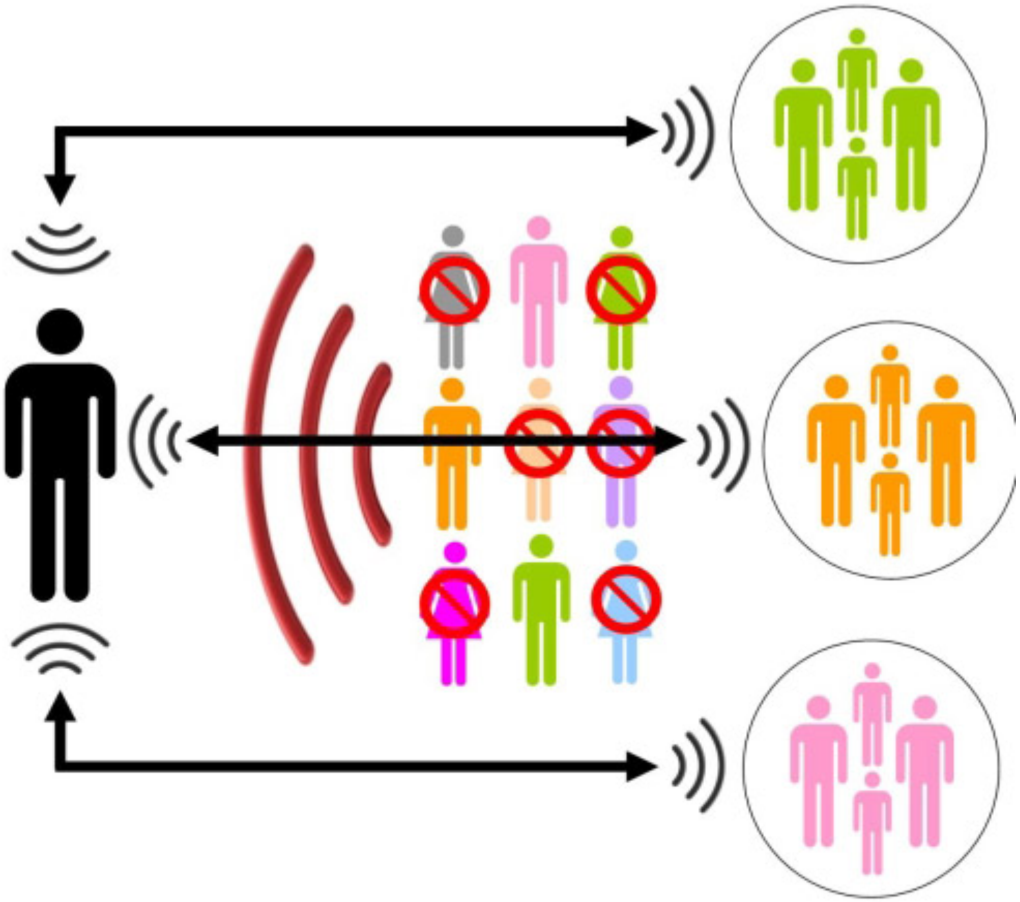
Content Related



User preference is an important driving factor for user behavior modeling.



Social Related



User behavior in social networks are highly dependent on social peers.

Context Related

Rich context info in social networks. How to couple them with behavioral modeling?

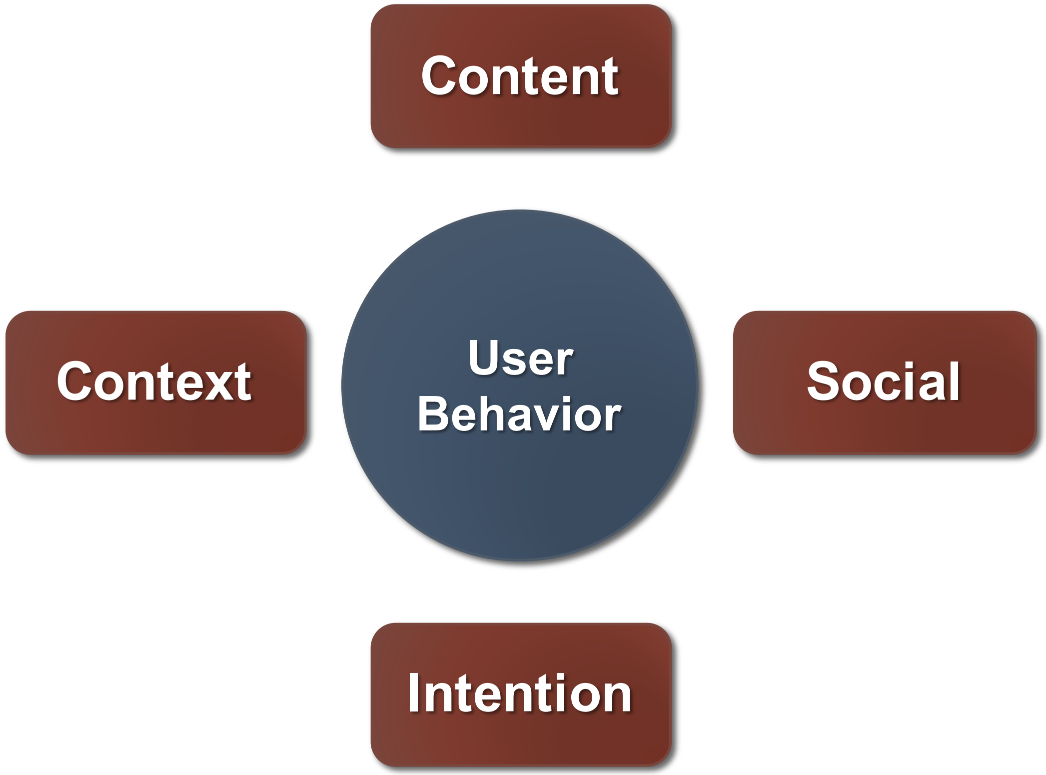
Intention Related

REWARDS	# TICKETS GIVEN	CONSEQUENCES	# TICKETS TAKEN AWAY
6 x 8 = 48 Extra Math	+5 YOU ARE A WINNER!	HITTING	-3 YOU ARE A WINNER!
Getting along WELL with others	+3 YOU ARE A WINNER!	BULLYING	-4 YOU ARE A WINNER!
Good Table Manners	+4 YOU ARE A WINNER!	TEASING	-1 YOU ARE A WINNER!
LOVE & RESPECT	+5 YOU ARE A WINNER!	LYING	-2 YOU ARE A WINNER!
Obeying the FIRST TIME	+3 YOU ARE A WINNER!	THROWING A FIT	-3 YOU ARE A WINNER!
Calm & Quiet in STORE	+3 YOU ARE A WINNER!	Ignoring Parents	-4 YOU ARE A WINNER!
Extra Reading	+2 YOU ARE A WINNER!	SCREAMING or YELLING	-1 YOU ARE A WINNER!
CLEANING up after PLAYING	+2 YOU ARE A WINNER!	BAD SPORT	-2 YOU ARE A WINNER!

A nontrivial part of user behaviors are from profitable and social purposes.

Intention can account for the behaviors that cannot be well interpreted by content, social and context.

Behavioral Modeling

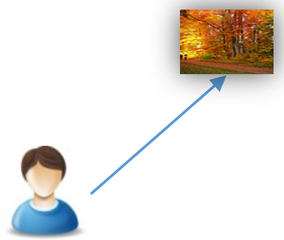


- Understand single factors
- Understand their couplings
- Unify them for behavior modeling and prediction

Granularities of User Behavior

User Behavior

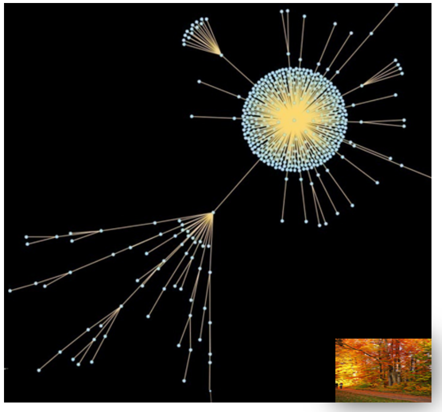
Microscopic



Personalized

Individual behavior analysis and prediction

Macroscopic



Global

Propagation analysis and prediction

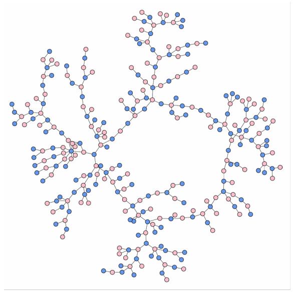
In This Tutorial

Micro

Natural behavior

Individual behavior

Information cascade



Intention

Unnatural/Suspicious behavior

Macro

Outline

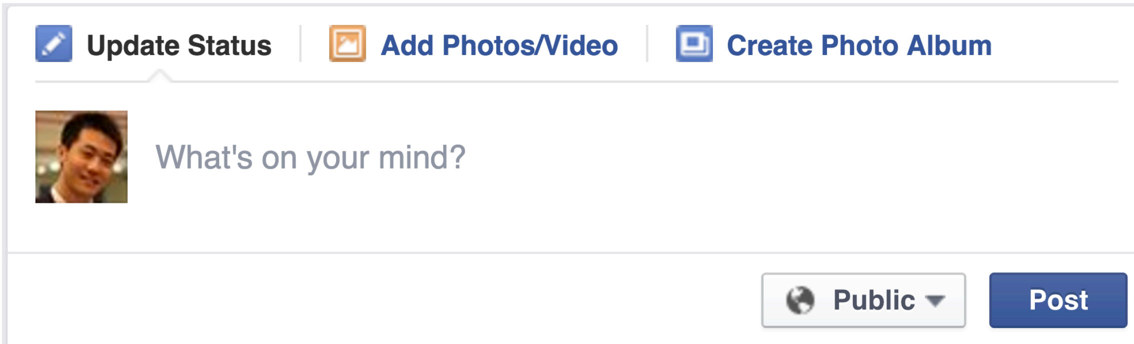
- **Prediction for natural behavior**
  - **Modeling individual behavior (MICRO)**
  - Modeling information cascade (MACRO)
- Detection for unnatural behavior
  - Suspicious behavior detection

Questions for Modeling Individual Behavior

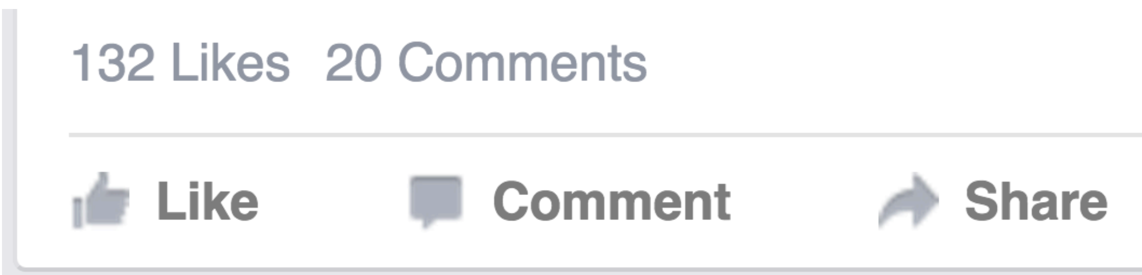
- What is individual behavior in social networks?
- Why should we study individual behavior?
- What are the state-of-the-art models?
  - Modeling behaviors and social relations
  - Modeling social contexts
  - Modeling spatiotemporal contexts
  - Modeling multiple domains in social networks

Individual Behavior: Facebook

- Post: What’s on your mind?

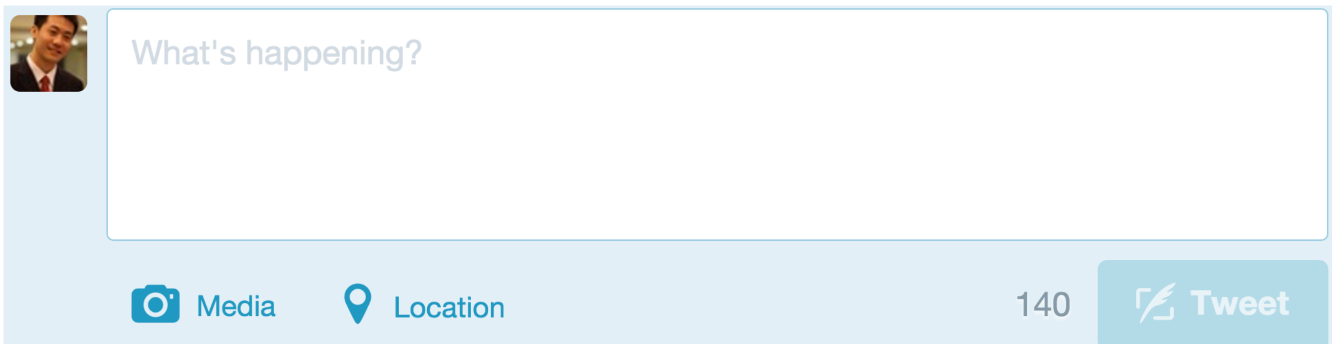


- Like, Comment, Share

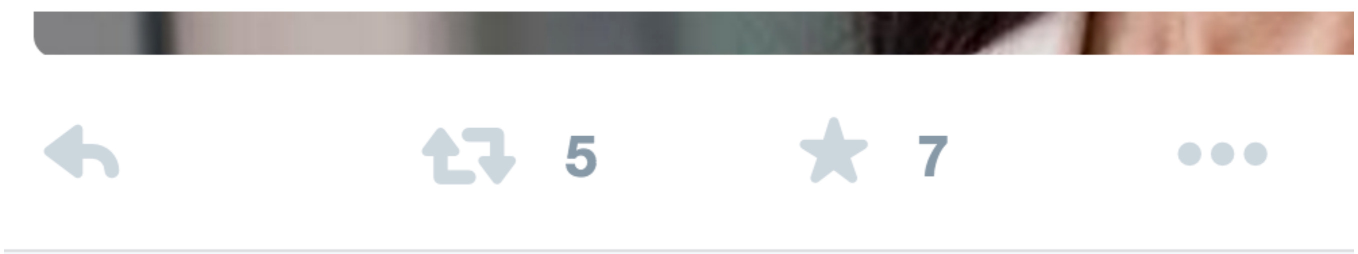


Individual Behavior: Twitter

- Tweet: What's happening?

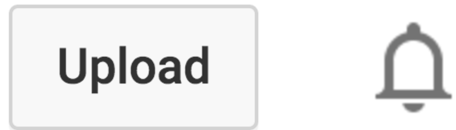


- Reply, Retweet, Favorite

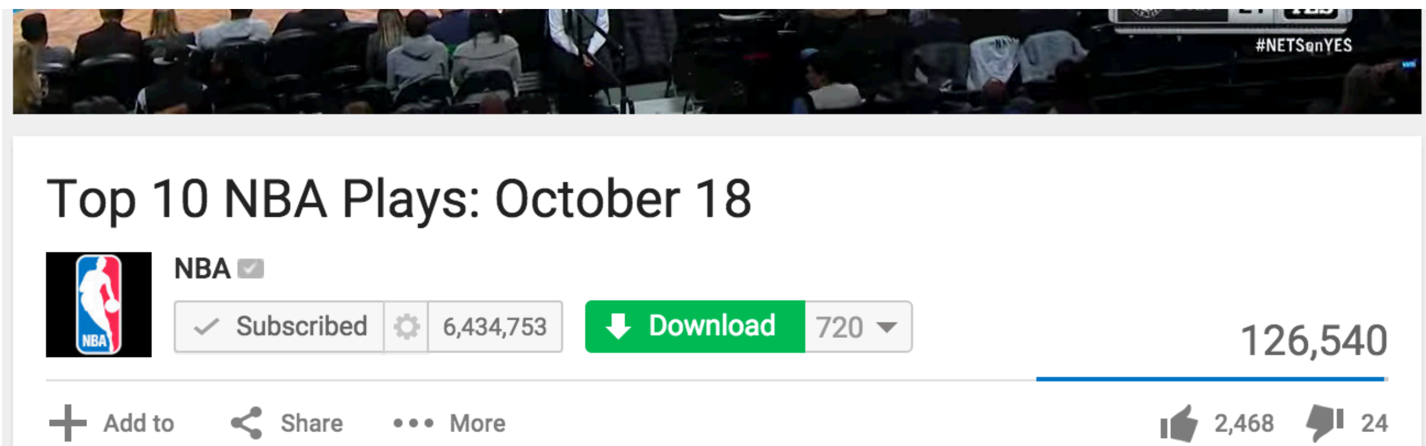


Individual Behavior: YouTube

- Upload

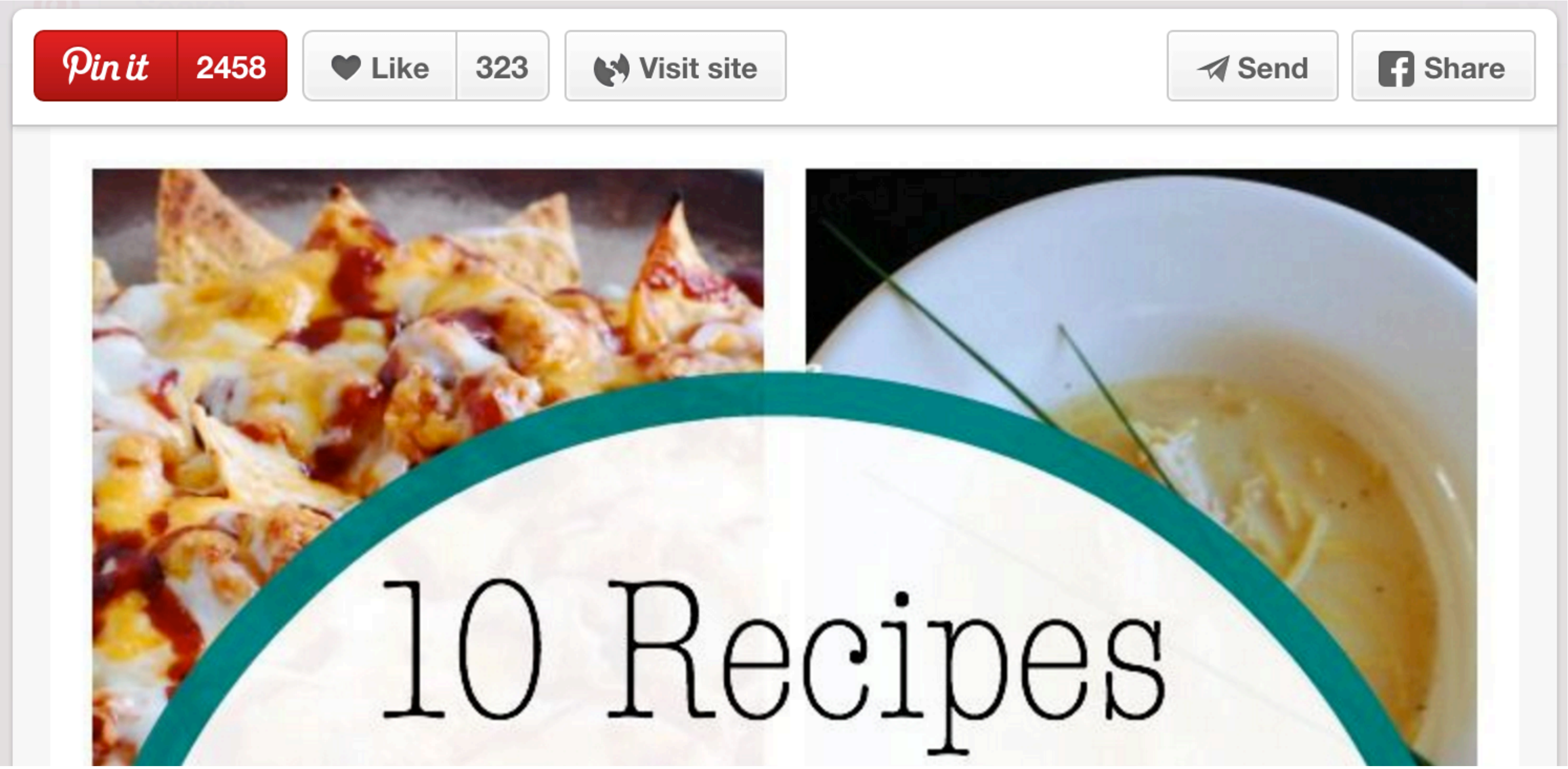


- Subscribe, Download, Add to, Share, Like, Dislike, Comment



Individual Behavior: Pinterest

- Pin it, Like, Visit, Send, Share



"Information Adoption" in Social Networks

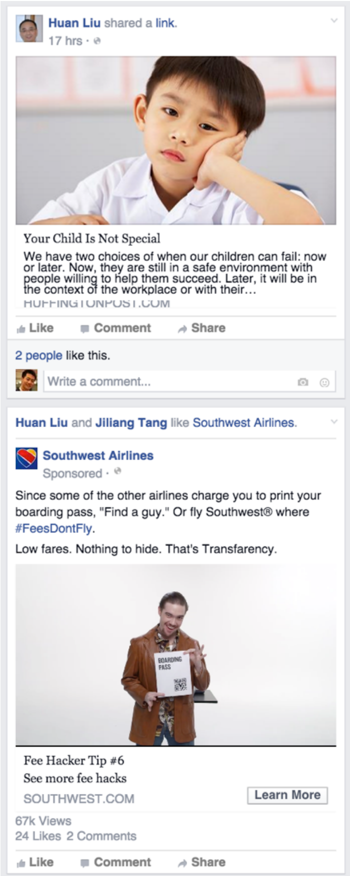
Like
Reply
Share
Favorite
Retweet
Comment
Subscribe
Download
Add to
Send
Pin it
Visit
……

Modeling Information Adoption Behavior

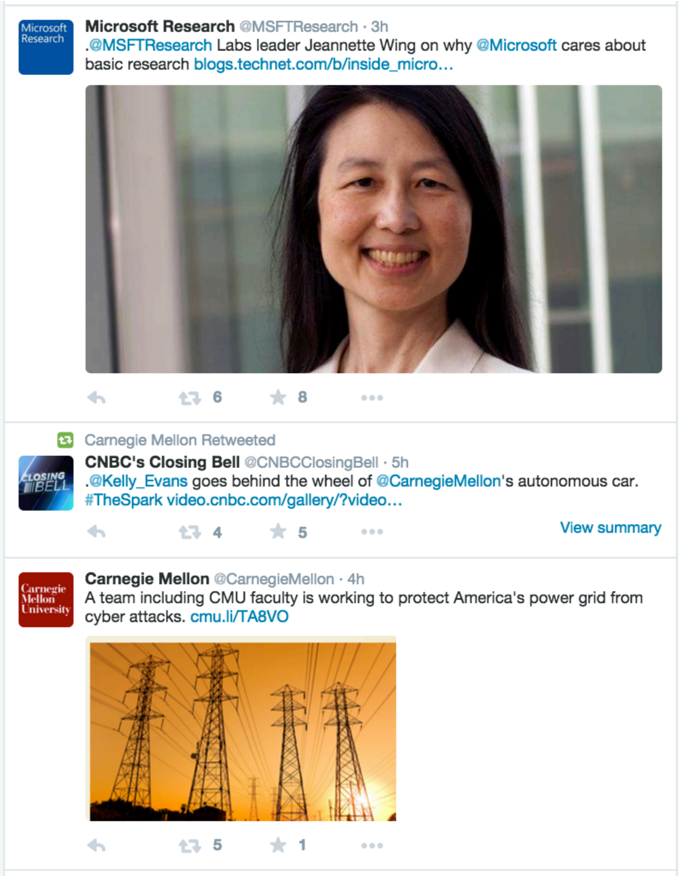
- Behavioral pattern discovery
- Behavior prediction in social networks
- Social recommendation

What is Social Recommendation?

Facebook



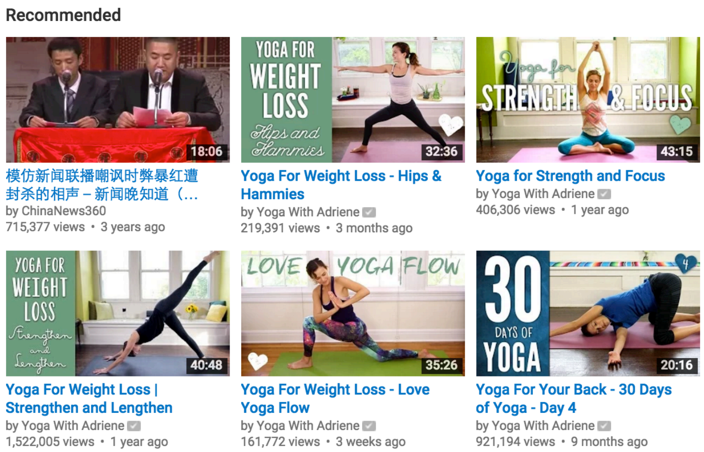
Twitter



News Feed Ranking

What is Social Recommendation?

YouTube



Pinterest

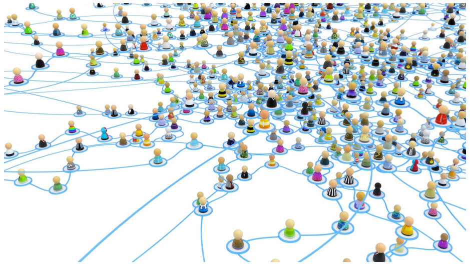
Social Multimedia Recommender Systems

What is Social Recommendation?

- "Social recommendation …, however, it has *no commonly accepted definition*."

What is Social Recommendation?

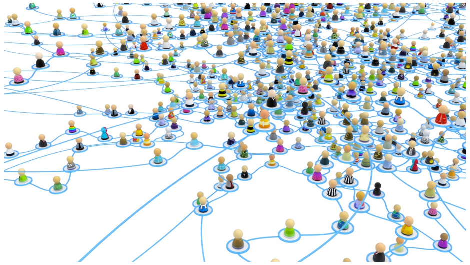
- “Social recommendation …, however, it has *no commonly accepted definition*.”
- “A narrow definition of social recommendation is *any recommendation with online social relations as an additional input*, i.e., augmenting an existing recommendation engine with *additional social signals*.”



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

What is Social Recommendation?

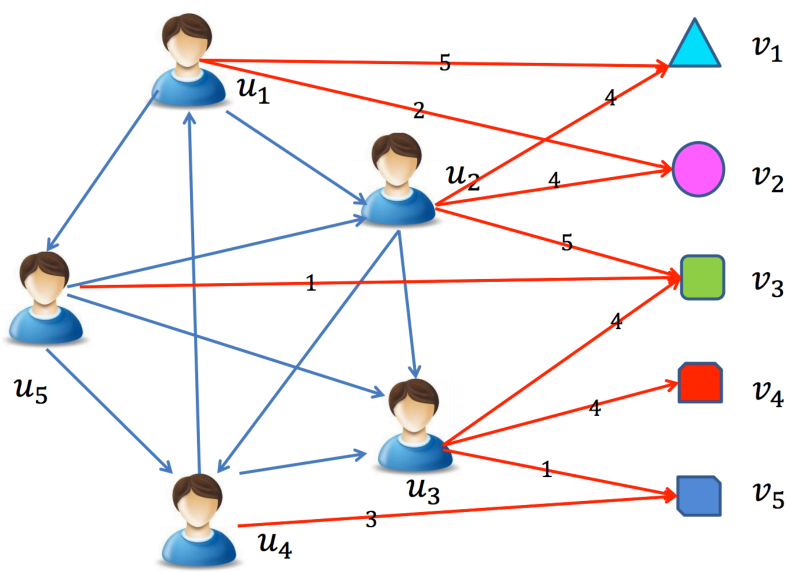
- “Social recommendation …, however, it has *no commonly accepted definition*.”
- “A narrow definition of social recommendation is *any recommendation with online social relations as an additional input*, i.e., augmenting an existing recommendation engine with *additional social signals*.”
- “Users’ preferences are likely to be similar to or influenced by their connected friends. …, social recommendation *leverages user correlations implied social relations* to improve the performance of recommendation.”



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

Traditional Recommender Systems

- Assumed that users are independent and identically distributed



	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	?	?

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

Traditional Recommender Systems

- Content-based recommender (e.g., TFIDF)
  - For textual information (e.g., news, documents)
  - *Limitation:* limited content analysis, over-specialization

Traditional Recommender Systems

- Content-based recommender (e.g., TFIDF)
  - 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

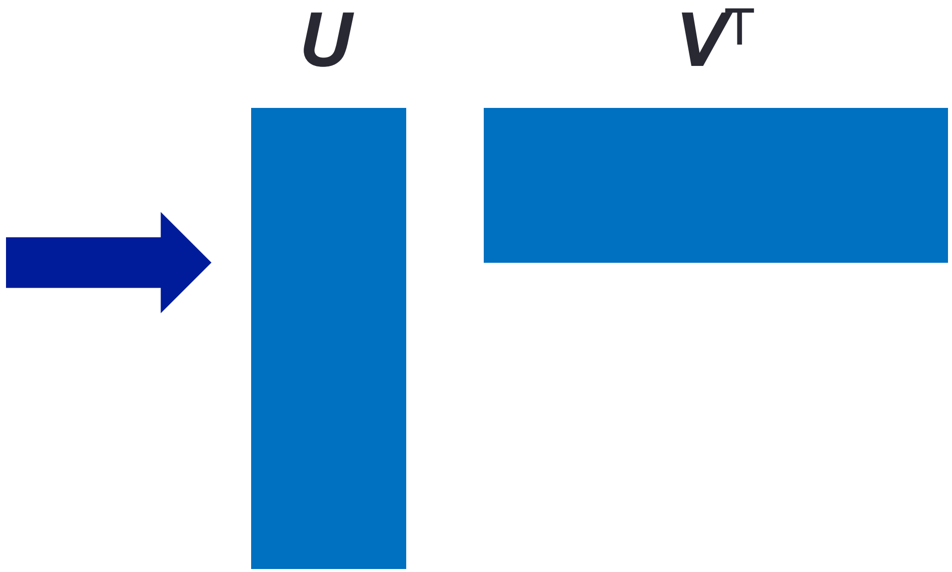
Tang et al. **Social Recommendation: A Review.** *Social Network Analysis and Mining*, 2013. Springer.

Matrix Factorization based CF (MF)

- 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 based CF (MF)

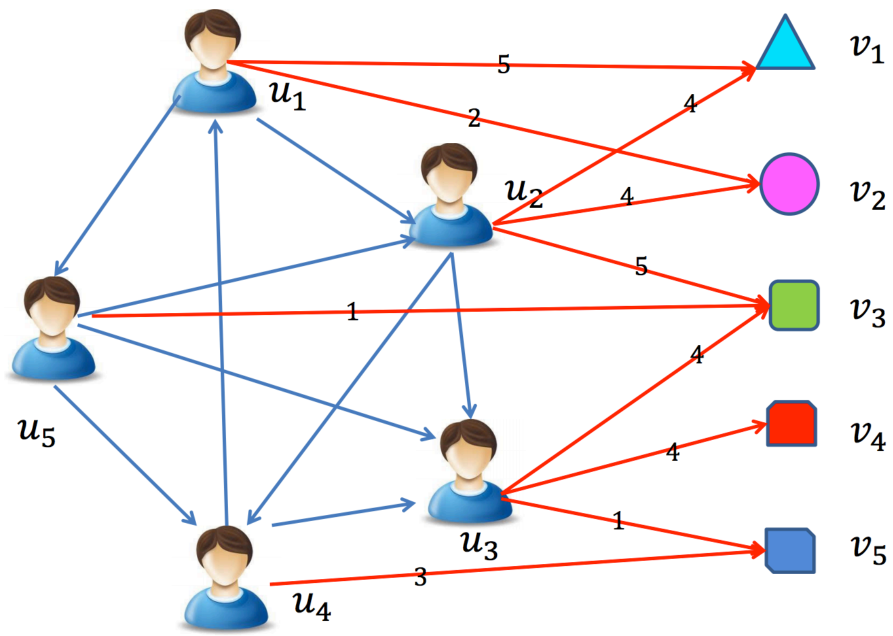
- Low-rank MF on the user-item rating matrix ***R***
- User preference vector ***U***
- Item characteristic vector ***V***
- Observed weight matrix ***W***

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

avoid over-fitting,
controlled by the parameter

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

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

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

rating
(user *i*, item *m*)

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*)

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.

Social Recommendation CF
= *Basic CF* + *Social Information Model*

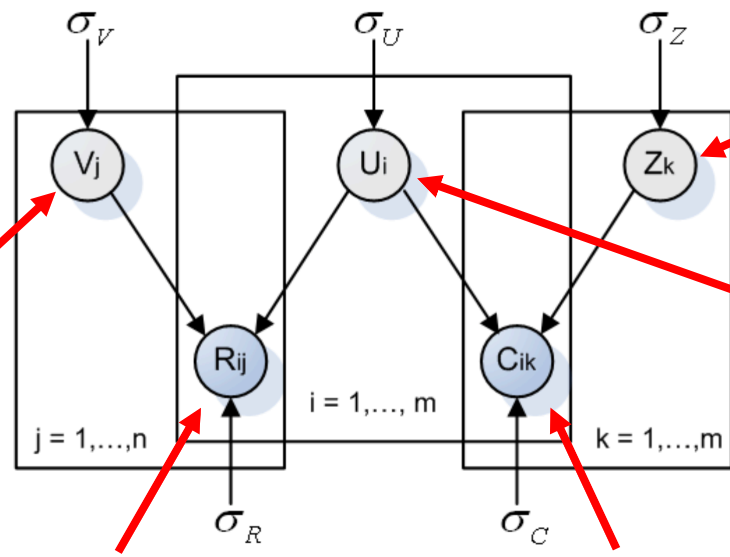
Tang et al. **Social Recommendation: A Review.** *Social Network Analysis and Mining*, 2013. Springer.

Model based Social Recommender

❖SoRec

factor vector

item vector



user vector

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

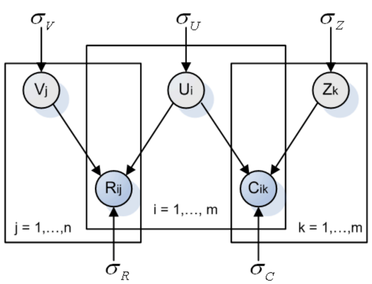
Ma et al. **Social Recommendation Using Probabilistic Matrix Factorization.** *CIKM*, 2008.
Improving Recommender Systems by Incorporating Social Contextual Information. *TIS*, 2011.

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

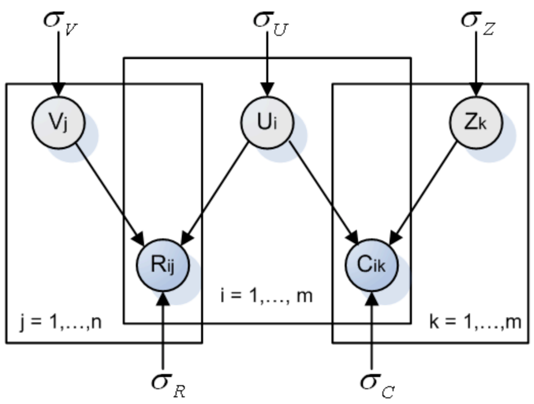
Observed

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

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



behavioral term

social term

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

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

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

regularization terms

Model based Social Recommender

❖SoRec

Gradient Descent Methods

deviate of Logistic function

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

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

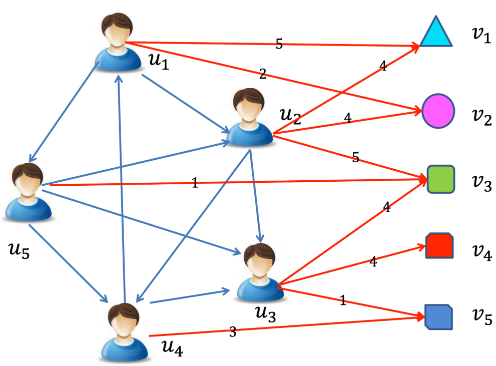
$$\frac{\partial \mathcal{L}}{\partial V_j} = \sum_{i=1}^{m} I_{ij}^{R} \underline{g'(U_i^T V_j)}(g(U_i^T V_j) - r_{ij})U_i + \lambda_V V_j,$$

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

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

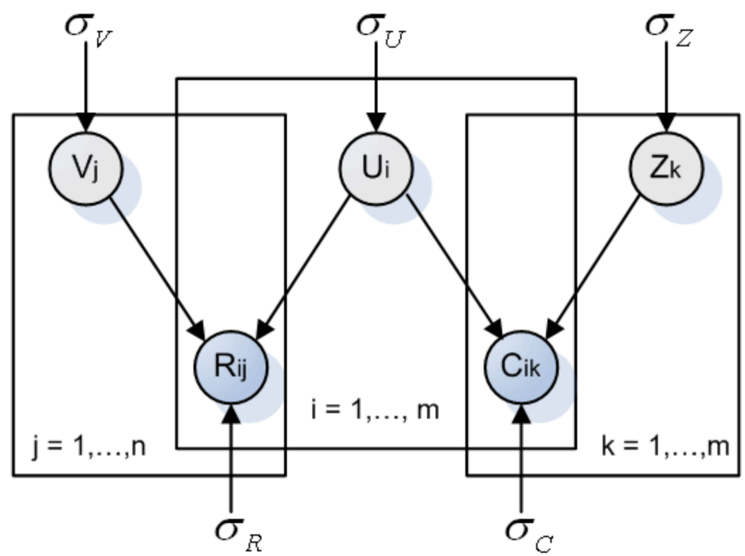
Model based Social Recommender

❖SoRec



	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

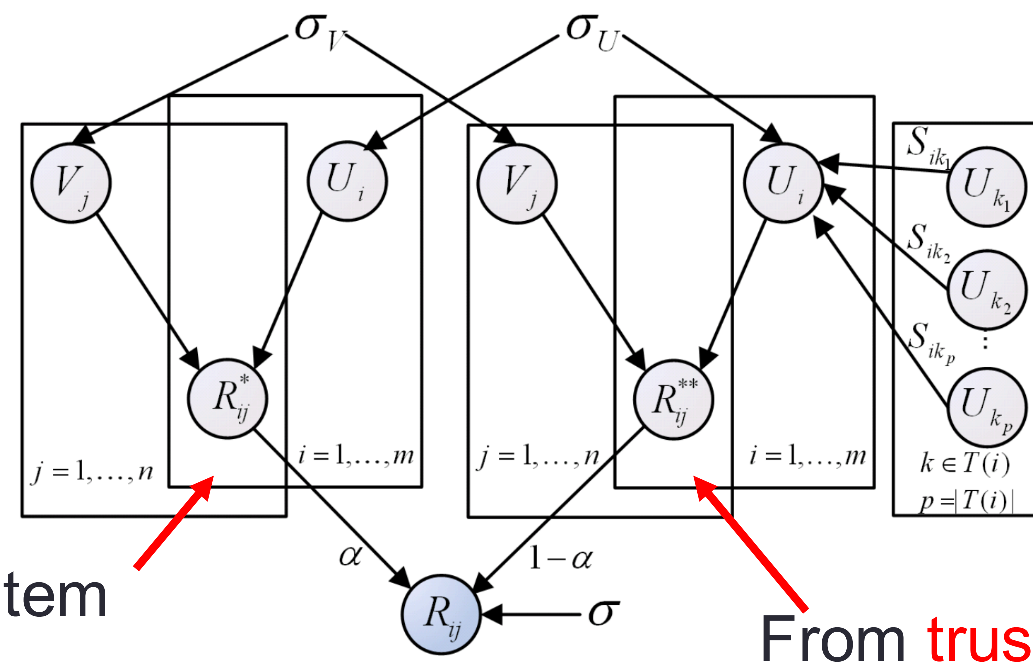
	v_1	v_2	v_3	v_4	v_5
u_1	5	?	2	?	?
u_2	4	4	5	?	?
u_3	?	?	4	4	1
u_4	?	?	?	?	3
u_5	?	?	1	?	?



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

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



From user-item matrix

Operation "plus"

From trusted friends

Model based Social Recommender

❖ "Social Trust" Ensemble

From user-item matrix

From trusted friends

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

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

Gradient Descent Methods

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

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

❖ SoReg

Average-based regularization:
Regularize with the average of friends' tastes

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

Information loss: Friends may have diverse tastes!!!

Model based Social Recommender

❖SoReg

Individual-based regularization:
Regularize with friends individually

$$\min_{U,V} \mathcal{L}_2(R,U,V) = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{ij}(R_{ij} - U_i^T V_j)^2$$

$$\boxed{+ \frac{\beta}{2}\sum_{i=1}^{m}\sum_{f\in\mathcal{F}^+(i)} Sim(i,f)\|U_i - U_f\|_F^2}$$

$$+ \lambda_1\|U\|_F^2 + \lambda_2\|V\|_F^2. \qquad (11)$$

Social Recommenders Before 2012

	Behavior	Social	Trust
SoRec [CIKM'08, TIS'11]	√	√	
"Social Trust" Ensemble [SIGIR'09, TIST'11]	√		√
SoReg [WSDM'11]	√	√	

$\underbrace{\hspace{6em}}$

<u>Social Contextual Information</u>

Social Contextual Information

- Twitter

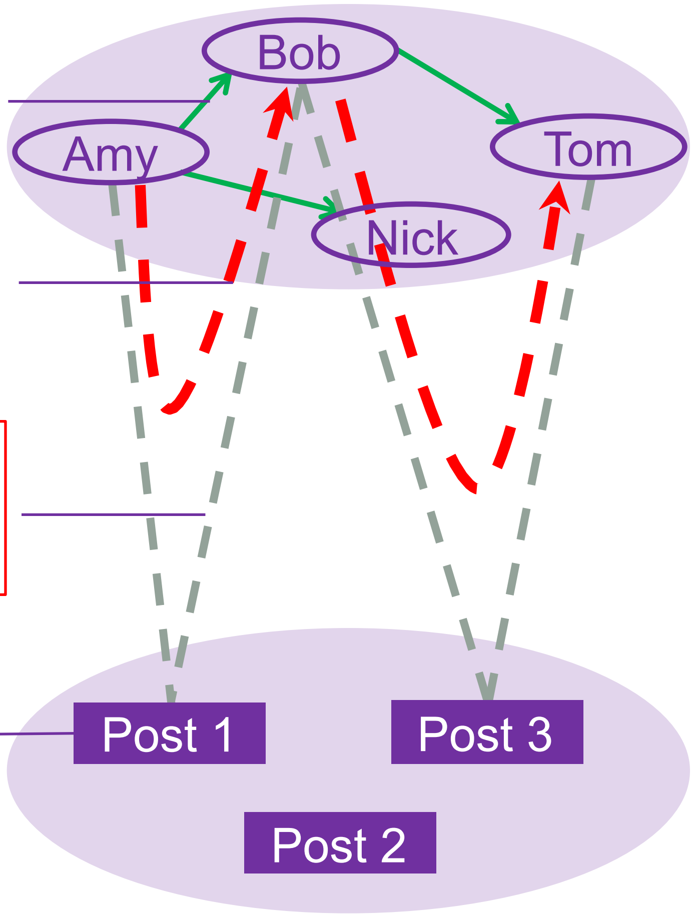
social relation = social

interaction frequency ≈ trust

retweeting/ rating = behavior

Additional signals

item content



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

Behavioral **Mechanism?**

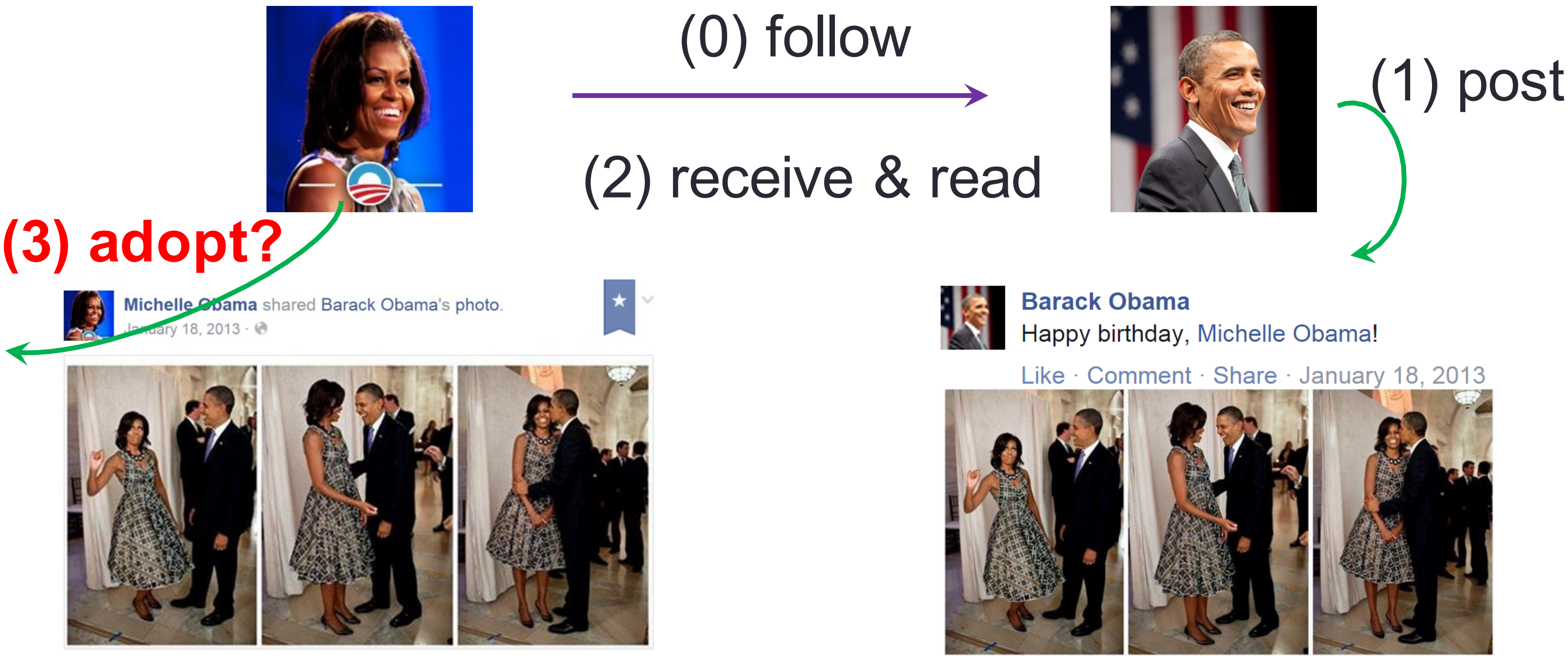
(0) follow

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

Information Adoption Behavior **Intention**

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

Information Adoption Behavior **Intention**



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

Information Adoption Behavior **Intention**

Birthday –
NOT politic issues!

(0) follow

(2) receive & read

(1) post

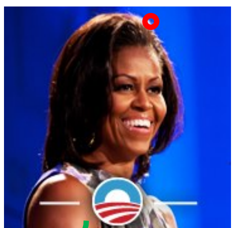
(3) adopt?

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

Information Adoption Behavior **Intention**

Birthday – NOT politic issues!

From my husband – NOT fans on Twitter!



(0) follow

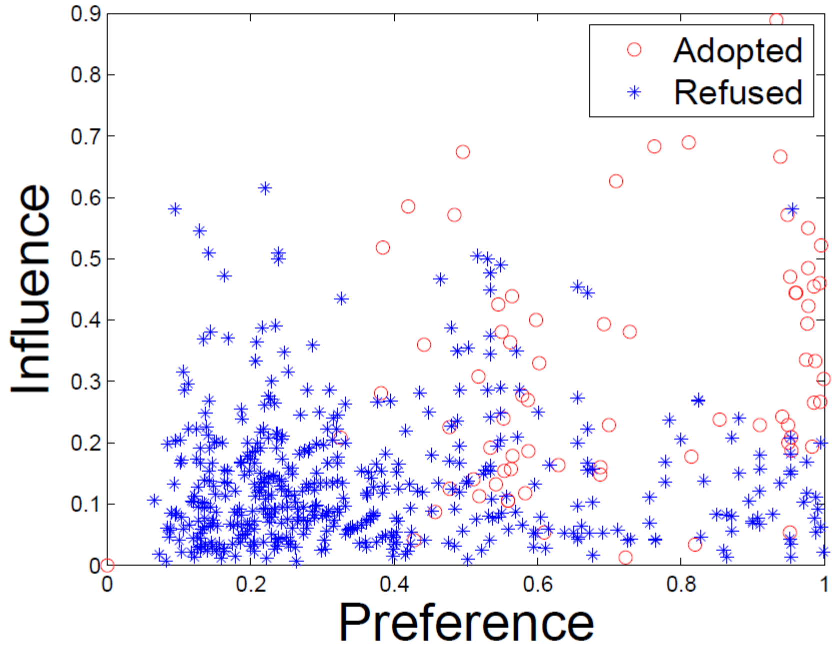
(2) receive & read

(1) post

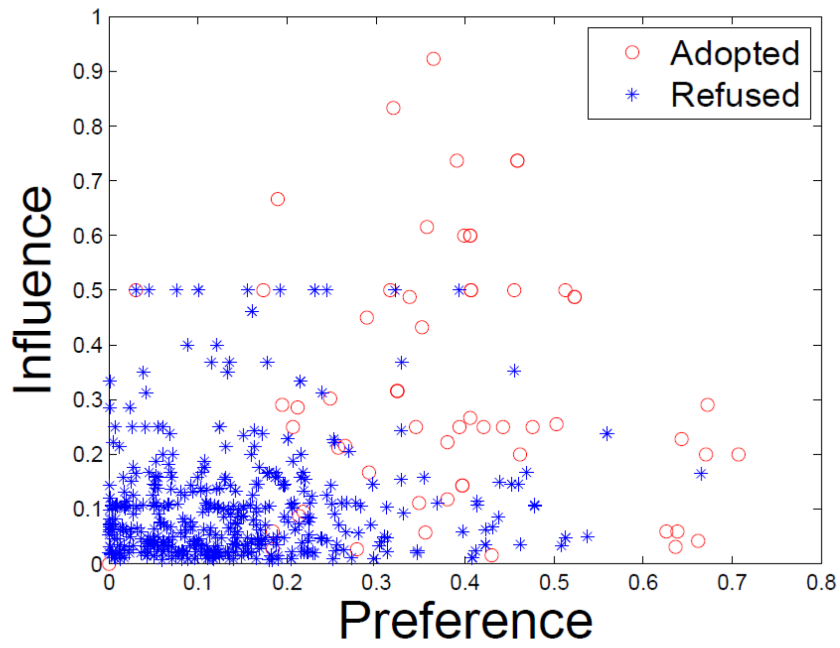
(3) adopt?

Social Contextual Factors

❖ Individual Preference & Interpersonal Influence



China's Facebook:
Renren



China's Twitter:
Tencent Weibo

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

From *Information* to *Factors*

Content:

item content

Behavior:

user-item interaction

Social:

social relation

Trust/interaction:

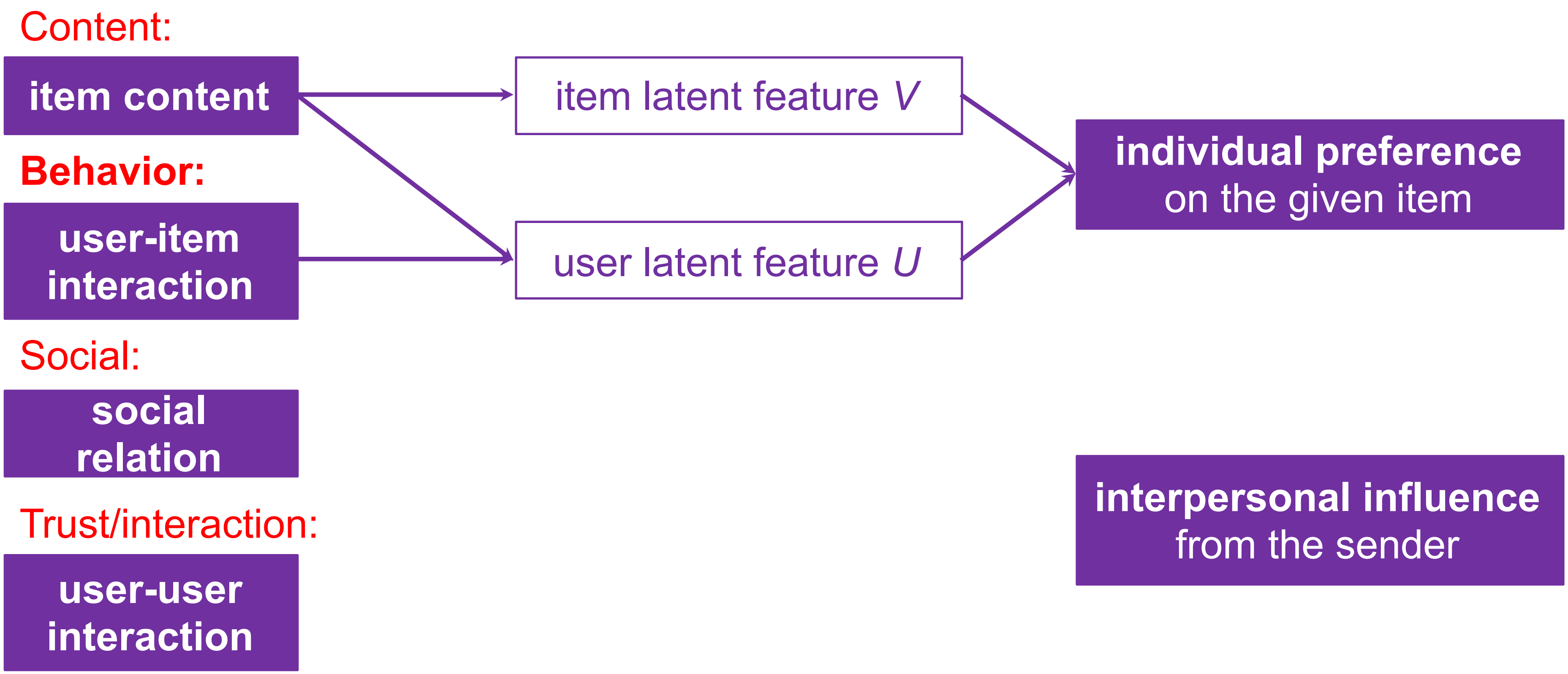
user-user interaction

individual preference on the given item

interpersonal influence from the sender

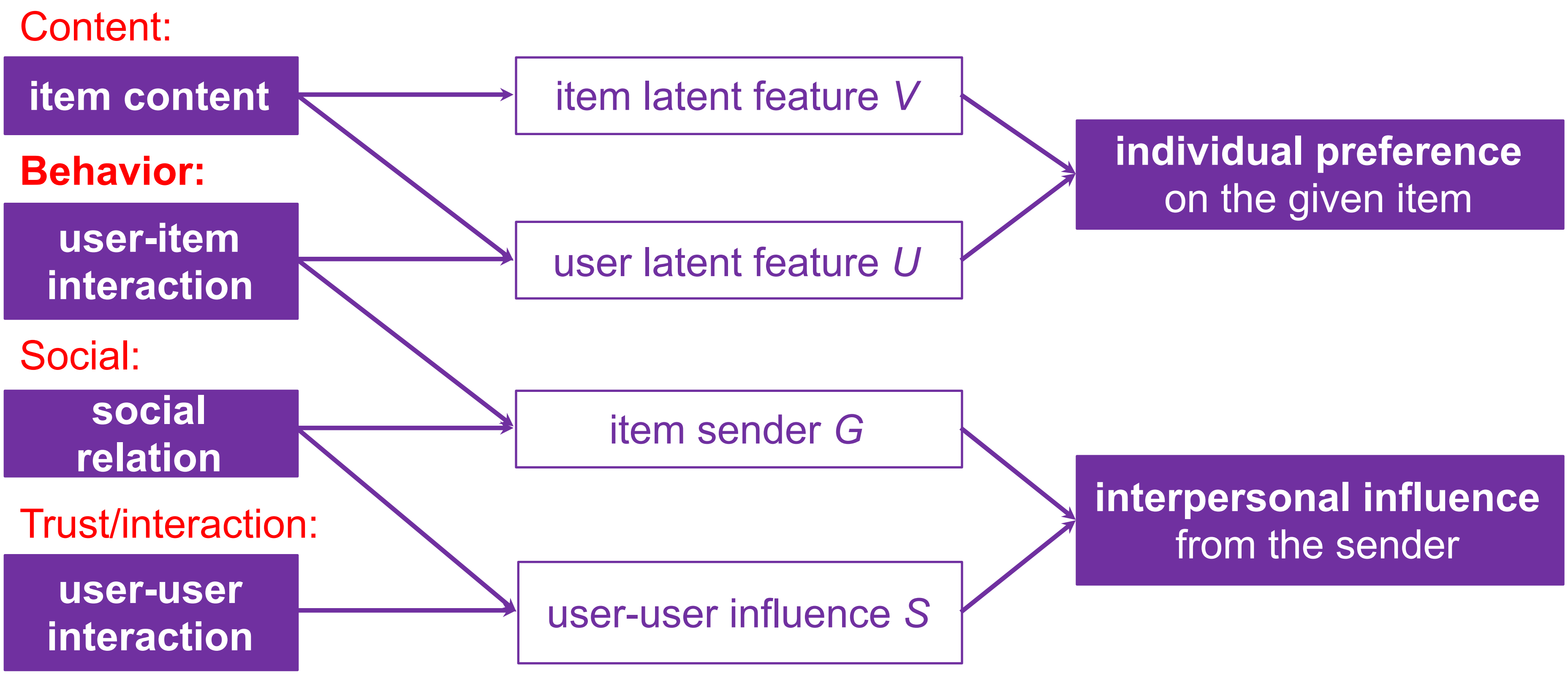
Jiang et al. **Social Contextual Recommendation.** *CIKM*, 2012. **Social Recommendation with Contextual Information.** *TKDE,* 2014.

From *Information* to *Factors*



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

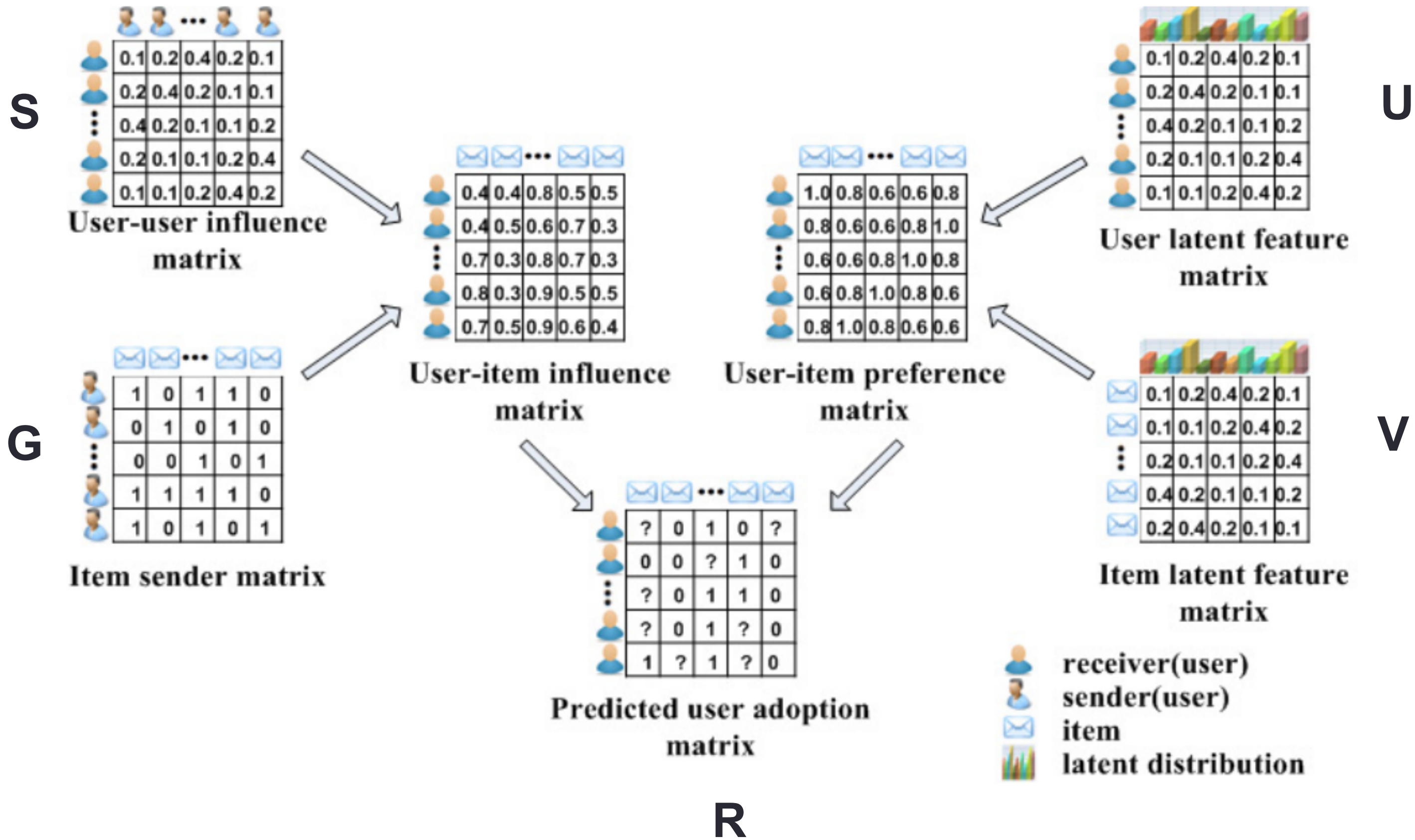
From *Information* to *Factors*



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

Social Contextual Recommendation

❖ContextMF



Social Contextual Recommendation

- ContextMF

behavior influence preference

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

behavior interaction frequency/trust

item content

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

social relation

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

Social Contextual Recommendation

- ContextMF

Gradient Descent Methods

$$\frac{\partial \mathcal{J}}{\partial \mathbf{S}} = 2\Big(-\mathbf{R}(\mathbf{G} \odot \mathbf{V}^\top \mathbf{U}) + (\mathbf{S}\mathbf{G}^\top \odot \mathbf{U}^\top \mathbf{V})\mathbf{G} + \gamma(\mathbf{S} - \mathbf{F}) + \delta \mathbf{S}\Big)$$

$$\frac{\partial \mathcal{J}}{\partial \mathbf{U}} = 2\Big(-\mathbf{V}\mathbf{R}^\top + \mathbf{V}(\mathbf{G}\mathbf{S}^\top \odot \mathbf{V}^\top \mathbf{U}) - 2\alpha \mathbf{U}\mathbf{W} + 2\alpha \mathbf{U}\mathbf{U}^\top \mathbf{U} + \eta \mathbf{U}\Big)$$

$$\frac{\partial \mathcal{J}}{\partial \mathbf{V}} = 2\Big(-\mathbf{U}\mathbf{R} + \mathbf{U}(\mathbf{S}\mathbf{G}^\top \odot \mathbf{U}^\top \mathbf{V}) - 2\beta \mathbf{V}\mathbf{C} + 2\beta \mathbf{V}\mathbf{V}^\top \mathbf{V} + \lambda \mathbf{V}\Big)$$

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

Social Contextual Recommendation

❖ContextMF

	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%

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**

Co-Predicting Behavior and Social Relations

❖LOCABAL

social relation

behavior

$$\mathcal{J} = \underline{\|\boxed{\mathbf{W} \odot} (\mathbf{R} - \mathbf{U}^\top \mathbf{V})}\|_F^2 + \alpha \underline{\|\boxed{\mathbf{T} \odot} (\mathbf{S} - \mathbf{U}^\top \mathbf{H} \mathbf{U})\|_F^2} + \lambda(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2 + \|\mathbf{H}\|_F^2) \quad (10)$$

Observed behavioral data

Observed social relation data

Tang et al. **Exploiting Local and Global Social Context for Recommendation.** *IJCAI*, 2013.

Co-Predicting Behavior and Social Relations

❖LOCABAL

Gradient Descent Methods

$$\frac{\partial \mathcal{J}}{\partial \mathbf{U}} = 2\big(- \mathbf{V}(\mathbf{W} \odot \mathbf{W} \odot \mathbf{R})^\top + \mathbf{V}(\mathbf{W} \odot \mathbf{W} \odot (\mathbf{U}^\top \mathbf{V}))^\top$$
$$- \alpha \mathbf{H}^\top \mathbf{U}(\mathbf{T} \odot \mathbf{S}) - \alpha \mathbf{H}\mathbf{U}(\mathbf{T} \odot \mathbf{S})^\top + \lambda \mathbf{U}$$
$$+ \alpha \mathbf{H}^\top \mathbf{U}(\mathbf{T} \odot (\mathbf{U}^\top \mathbf{H}\mathbf{U})) + \alpha \mathbf{H}\mathbf{U}(\mathbf{T} \odot (\mathbf{U}^\top \mathbf{H}\mathbf{U}))^\top\big),$$
$$\frac{\partial \mathcal{J}}{\partial \mathbf{V}} = 2\big(- \mathbf{U}(\mathbf{W} \odot \mathbf{W} \odot \mathbf{R})$$
$$+ \mathbf{U}(\mathbf{W} \odot \mathbf{W} \odot (\mathbf{U}^\top \mathbf{V})) + \lambda \mathbf{V}\big),$$
$$\frac{\partial \mathcal{J}}{\partial \mathbf{H}} = 2\big(- \lambda \mathbf{U}(\mathbf{T} \odot \mathbf{S})\mathbf{U}^\top$$
$$+ \alpha \mathbf{U}(\mathbf{T} \odot (\mathbf{U}^\top \mathbf{H}\mathbf{U}))\mathbf{U}^\top + \lambda \mathbf{H}\big) \qquad (11)$$

Co-Predicting Behavior and Social Relations

❖LOCABAL

Datasets	Training Set	Metrics	Algorithms			
			MF	**SoRec**	**SoReg**	LOCABAL
Ciao	50%	MAE	0.9927	0.9619	0.9552	0.9356
		RMSE	1.1742	1.1375	1.1291	1.1088
	70%	MAE	0.9715	0.9446	0.9328	0.9234
		RMSE	1.1478	1.1140	1.1097	1.0861
	90%	MAE	0.9614	0.9433	0.9232	0.9076
		RMSE	1.1384	1.1028	1.0999	1.0758
Epinions	50%	MAE	0.9935	0.9574	0.9383	0.9237
		RMSE	1.1922	1.1581	1.1479	1.1276
	70%	MAE	0.9701	0.9480	0.9296	0.9088
		RMSE	1.1833	1.1482	1.1277	1.1079
	90%	MAE	0.9687	0.9397	0.9188	0.8981
		RMSE	1.1791	1.1387	1.1170	1.1000

Tang et al. **Exploiting Local and Global Social Context for Recommendation.** *IJCAI*, 2013.

Negative Experiences in Social Recommender

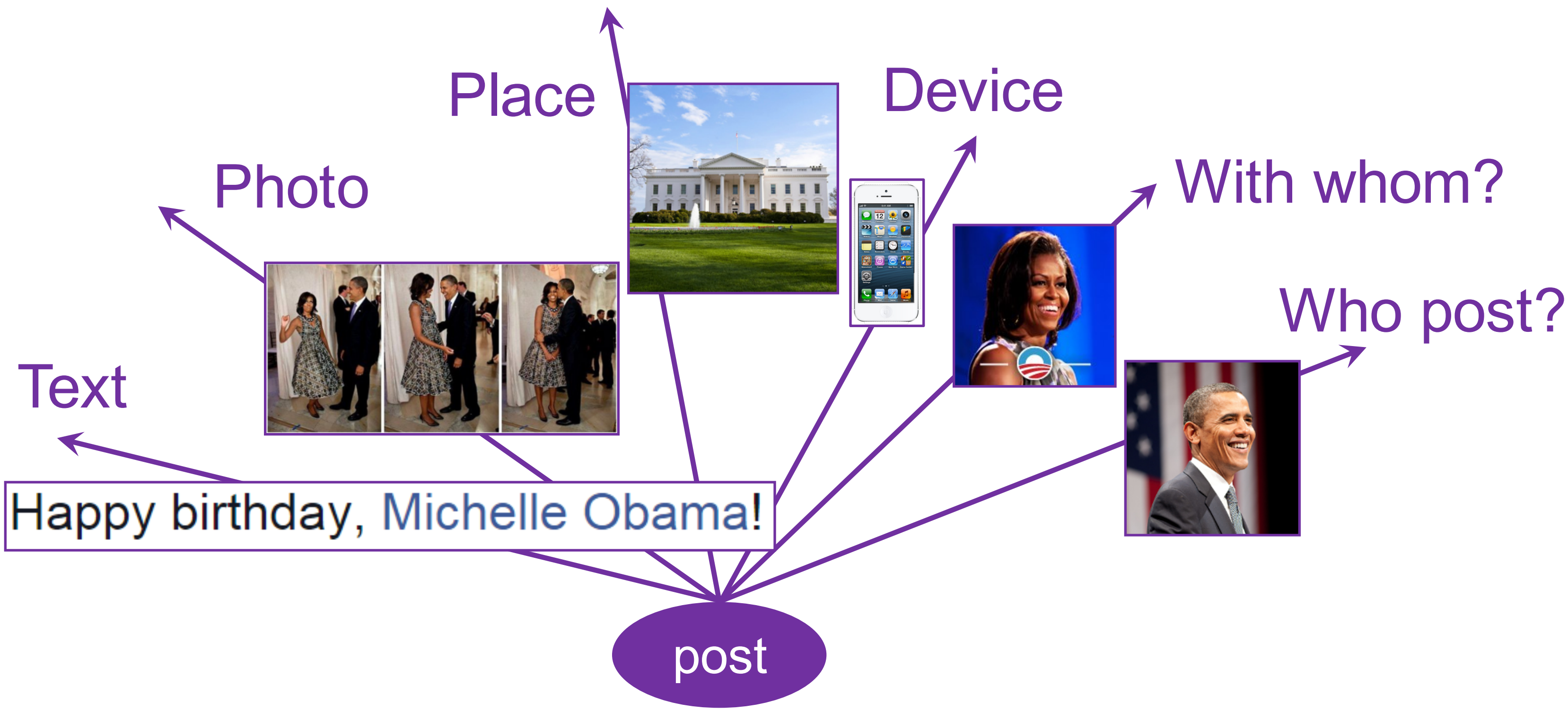
- Data sparsity problem: sparse matrix.
- Social relation are too noisy and may have a negative impact on recommender systems.
- It is difficult for social recommenders to improve recommendation performance for cold-start users.
- Different types of social relations have different effects on social recommender systems.

Negative Experiences in Social Recommender

- Data sparsity problem: sparse matrix.
- Social relation are too noisy and may have a negative impact on recommender systems.
- It is difficult for social recommenders to improve recommendation performance for cold-start users.
- Different types of social relations have different effects on social recommender systems.

Future work?
Data integration and causal analysis

Besides Social Context: Spatial Context



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

Besides Social Context: Temporal Context

Photo

Text

Happy birthday, Michelle Obama!

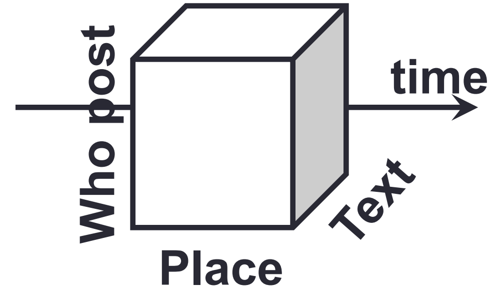
post → post → time

Jan. 18
Birthday party
@ White house

March 17
St. Patrick's Day
@ Bar

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

Modeling Spatiotemporal Contexts



Behavior

Spatial info: Multi-faceted

Temporal info: Dynamic

Tasks

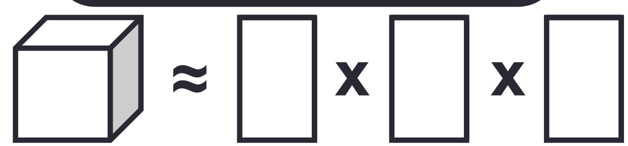
Behavior modeling

Pattern discovery

Behavior prediction

Tensor sequence

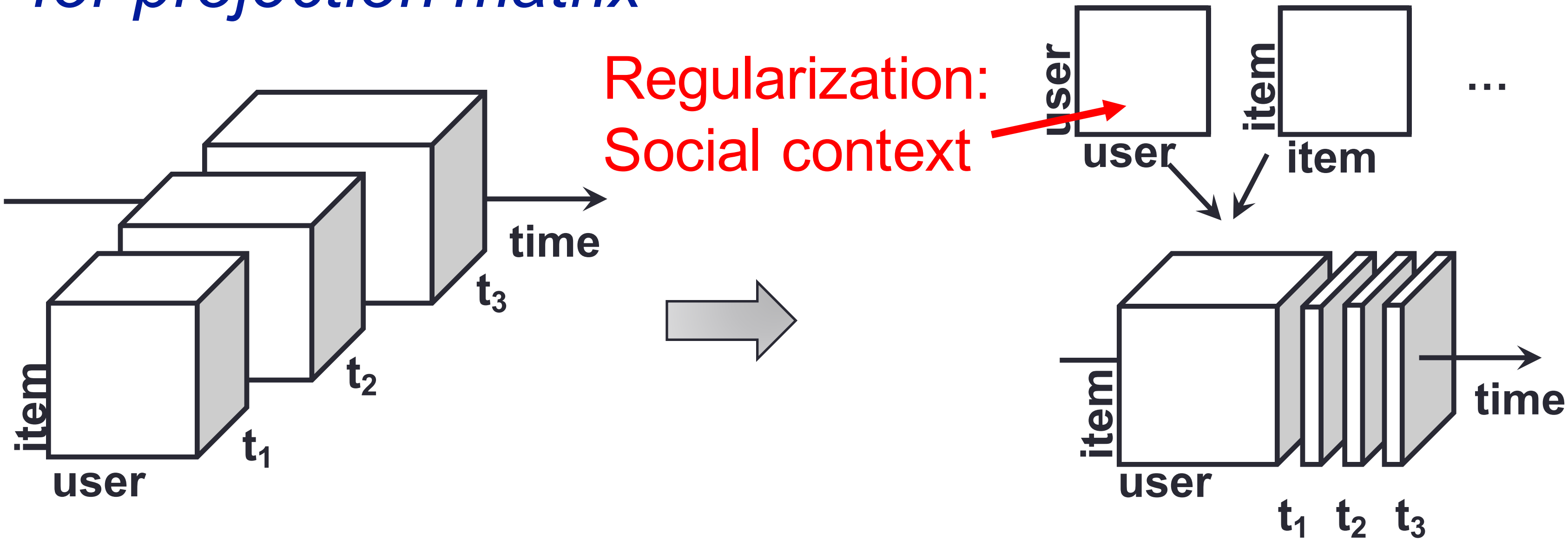
Decomposition & Completion



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

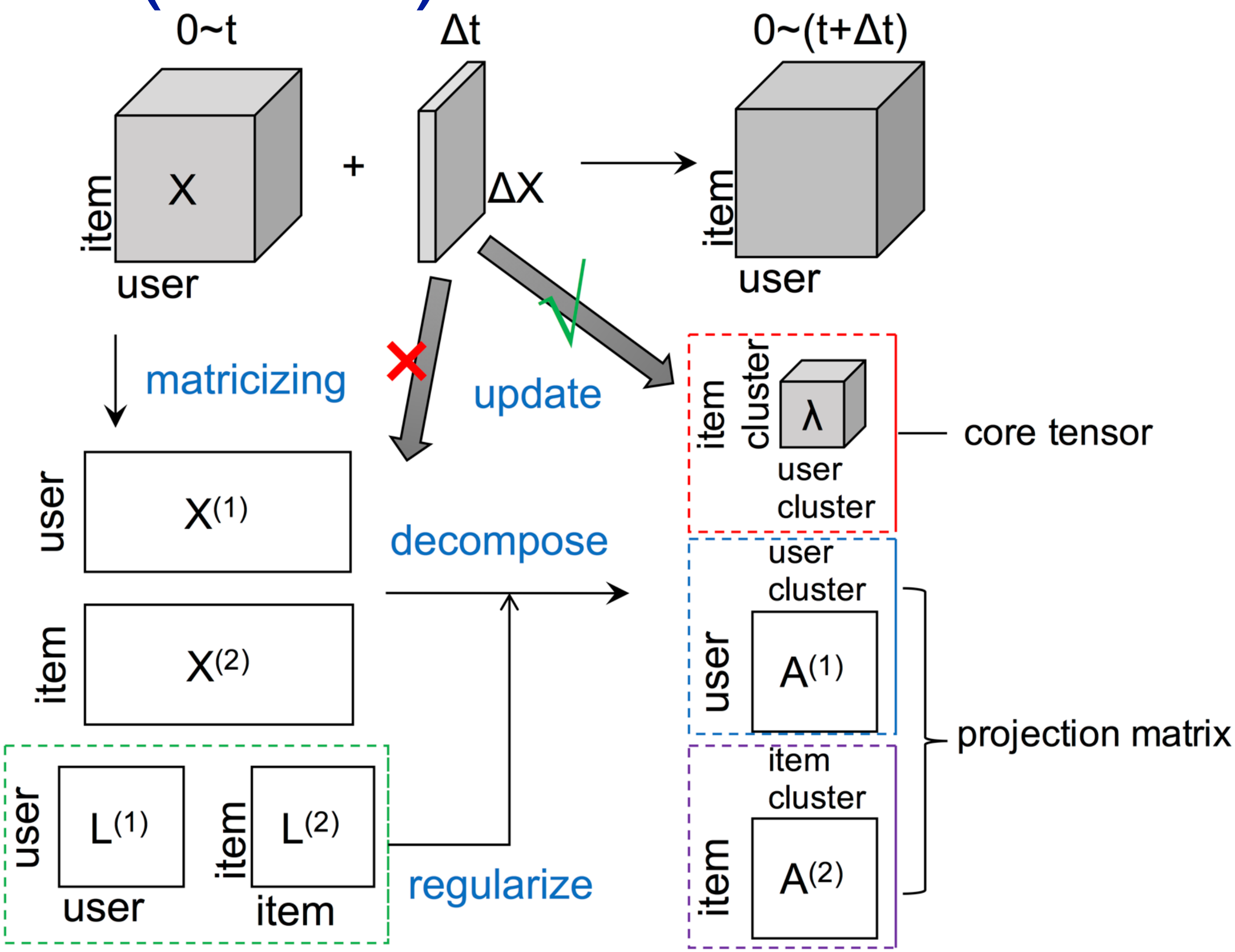
Challenges: Sparsity and Complexity

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



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

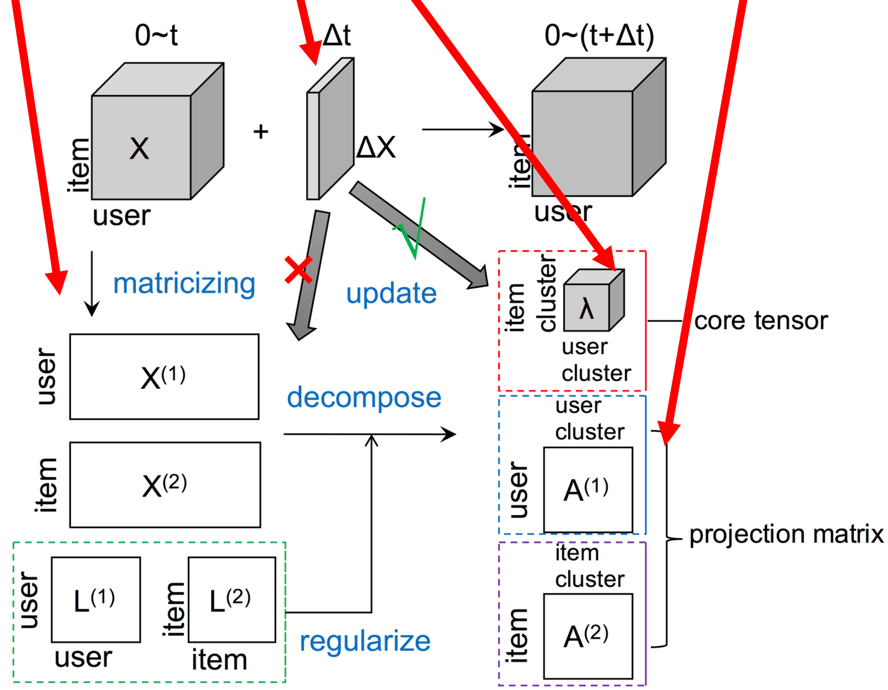
Flexible Evolutionary Multi-faceted Analysis (FEMA)



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

Tensor Perturbation Theory

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



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

FEMA Algorithm

Approximation

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

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

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

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

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

and compute

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

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

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

and compute

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

end for

end for

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

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

Bound Guarantee

core tensor

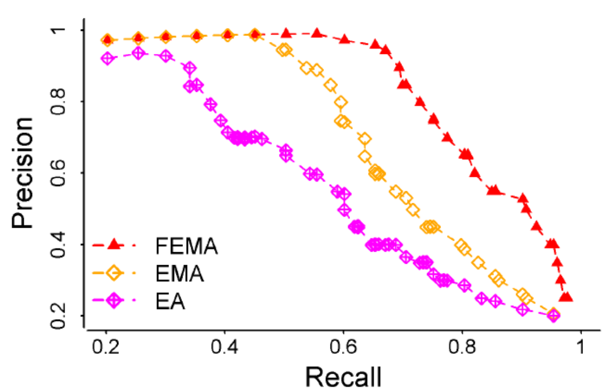
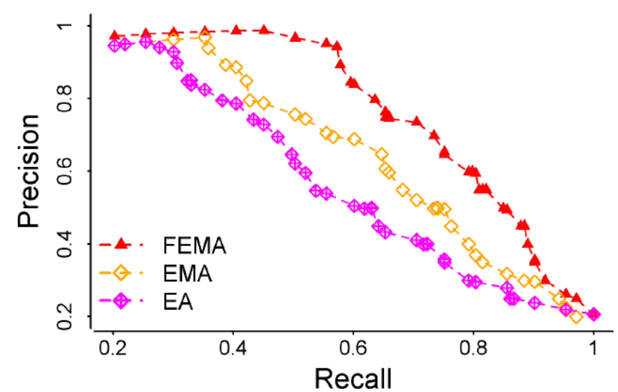
$$|\Delta\lambda_i^{(m)}| \leq 2(\lambda^{\max}_{\mathbf{X}^{(m)\top}\mathbf{X}^{(m)}})^{\frac{1}{2}}\|\Delta\mathbf{X}^{(m)}\|_2$$

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

projection matrix

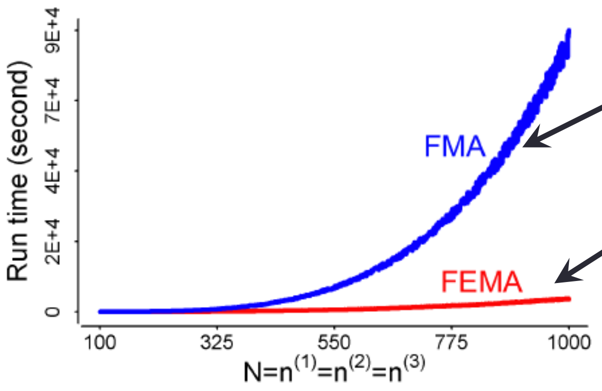
Jiang et al. **Flexible Evolutionary Multi-faceted Analysis for Dynamic Behavioral Pattern Discovery.** *KDD*, 2014.

Individual Behavior Prediction with FEMA

	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				

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

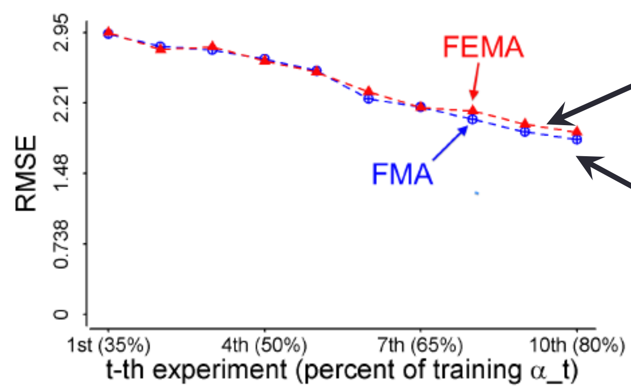
Individual Behavior Prediction with FEMA



Time vs Num. objects N

Re-decomposition:
Re-compute projection matrices

Evolutionary analysis:
Use ΔX to update λ and a



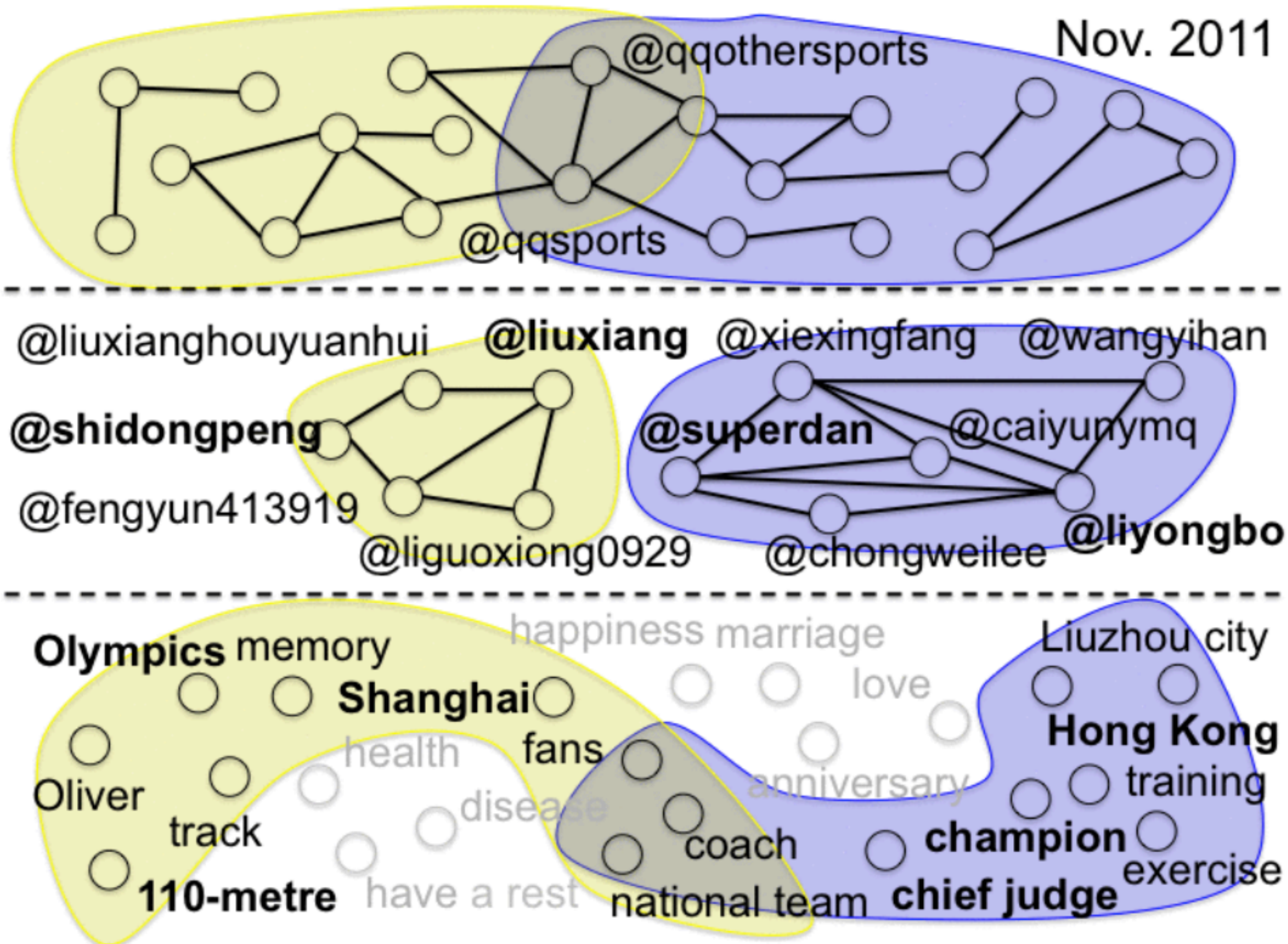
The loss is small.

Evolutionary analysis:
Use ΔX to update λ and a

Re-decomposition:
Re-compute projection matrices

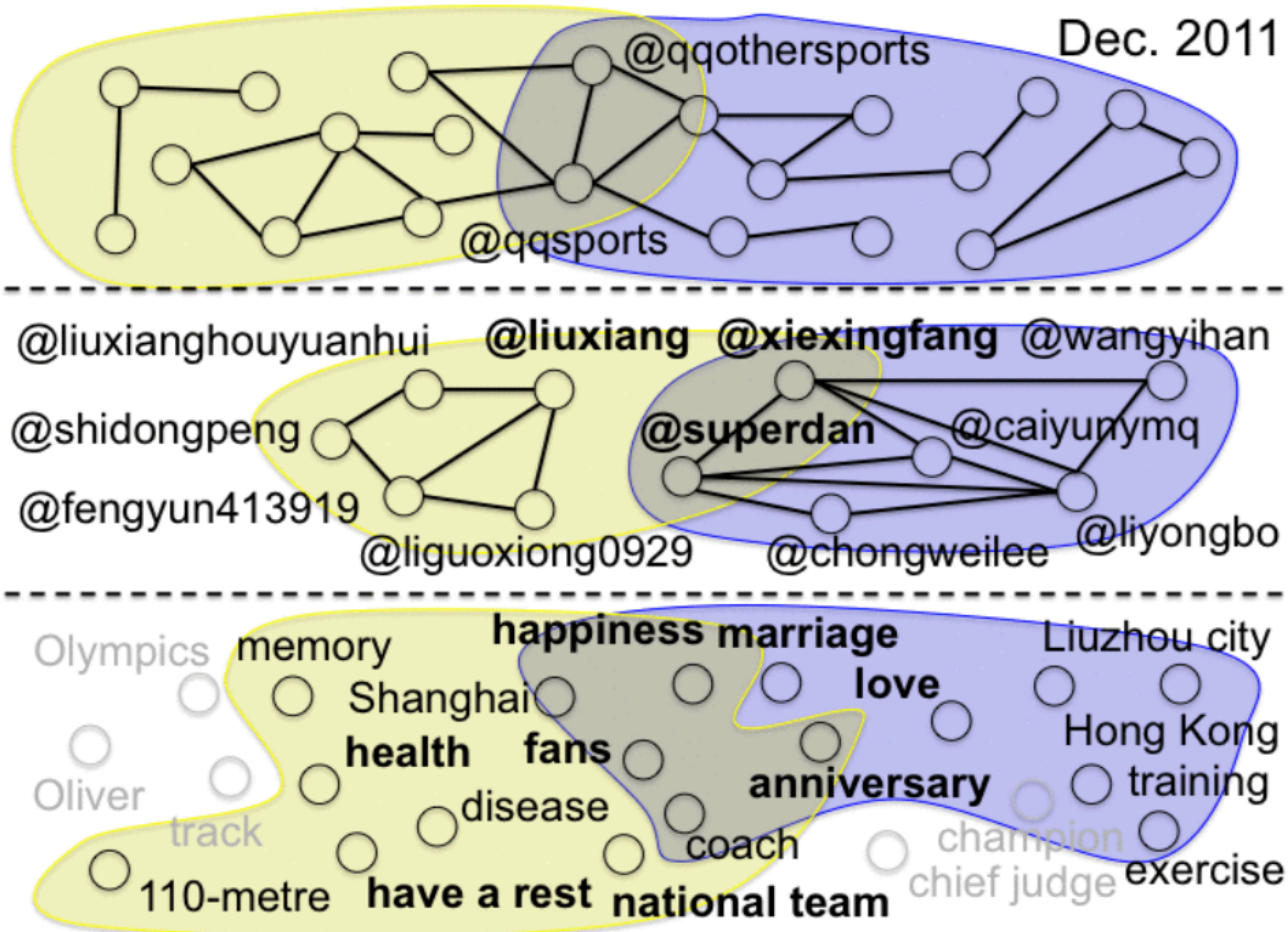
Jiang et al. **Flexible Evolutionary Multi-faceted Analysis for Dynamic Behavioral Pattern Discovery.** *KDD*, 2014.

Behavioral Pattern: Fan@Idol#Word



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

Behavioral Pattern: Fan@Idol#Word



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

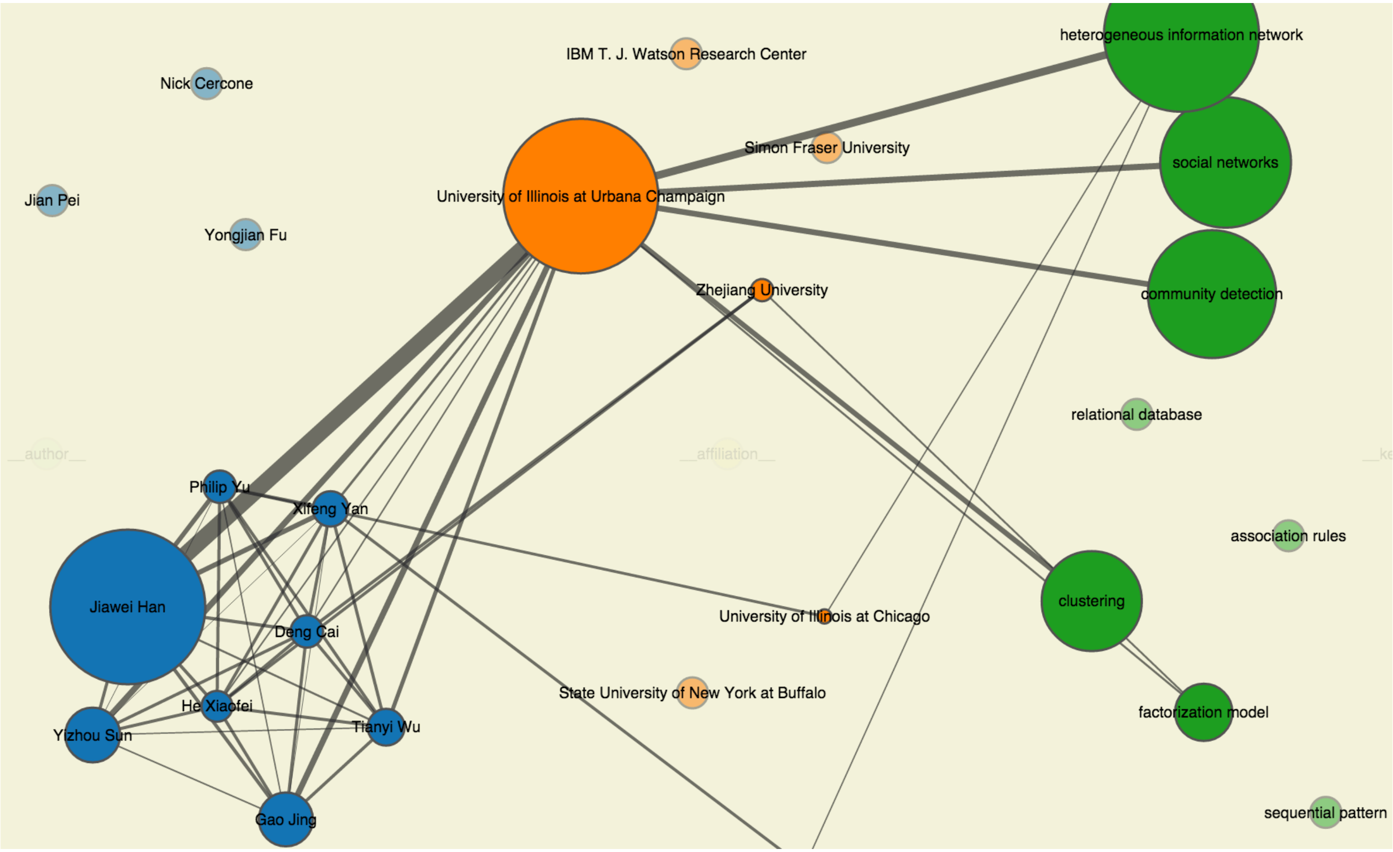
Demo 1: Fan@Idol#Word

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

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

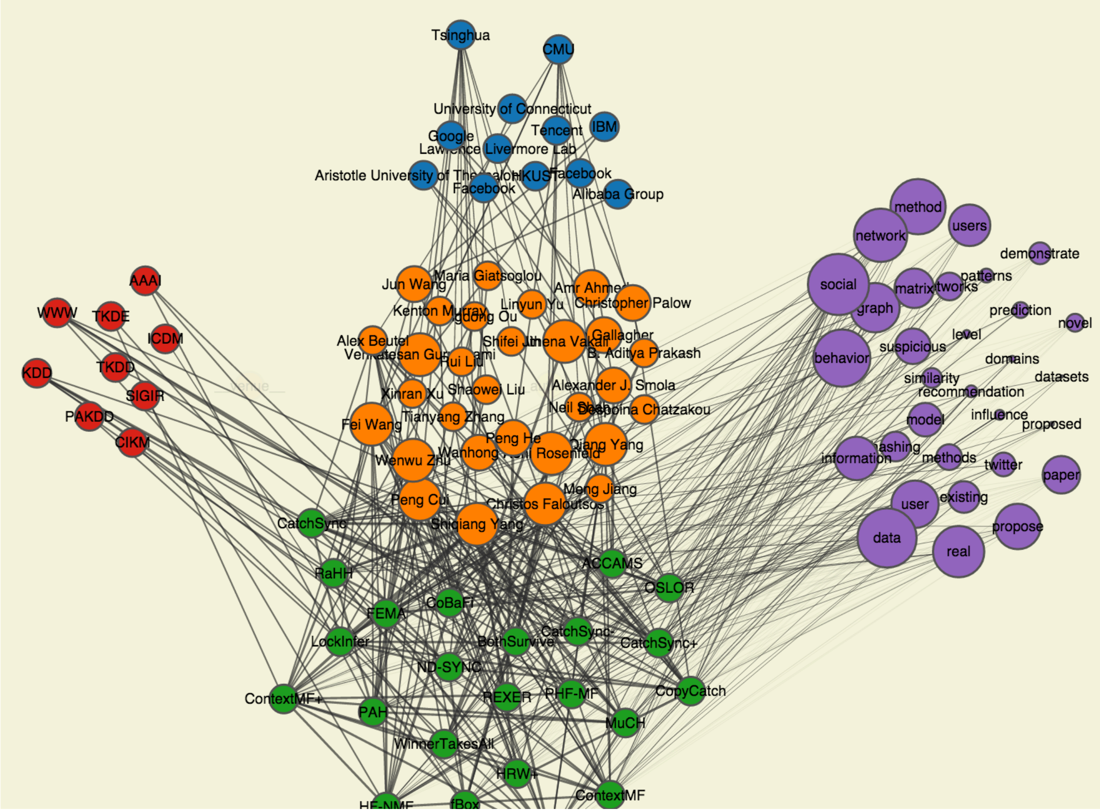
Demo 2: Author@Affiliation#Keyword

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



Demo 3: Author@Affiliation$Paper&Venue#Keyword

- http://www.meng-jiang.com/demos/hindblp/



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

Modeling social contexts, spatio-temporal contexts: Amazing!

…however…

Modeling social contexts, spatio-temporal contexts: Amazing!

…however…

Sparsity

in one domain…

Besides Contexts: Multiple Domains

- Post

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

Besides Contexts: Multiple Domains

- Image

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

Besides Contexts: Multiple Domains

- Video

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

Besides Contexts: Multiple Domains

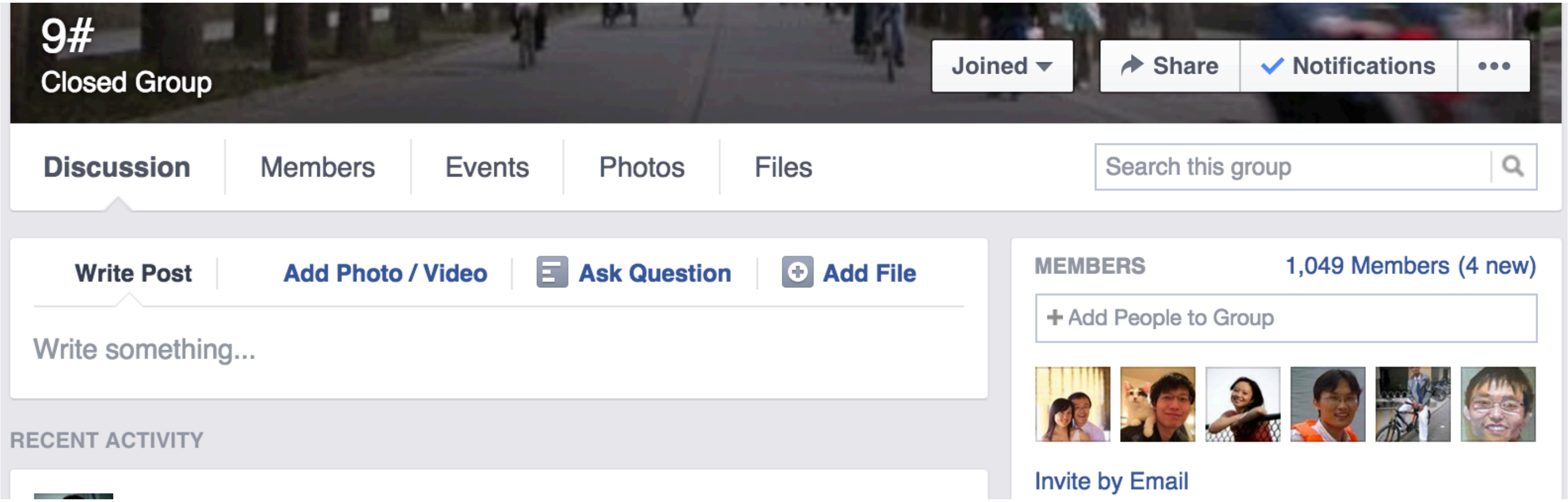
- Social label

Religious Views	Christian
Interests	Basketball, writing, spending time w/ kids
Favorite Music	Miles Davis, John Coltrane, Bob Dylan, Stevie Wonder, Johann Sebastian Bach (cello suites), and The Fugees
Favorite Movies	Casablanca, Godfather I & II, Lawrence of Arabia and One Flew Over the Cuckoo's Nest
Favorite TV Shows	Sportscenter
Favorite Quotations	"The Arc of the moral universe is long, but it bends towards justice." (MLK)

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

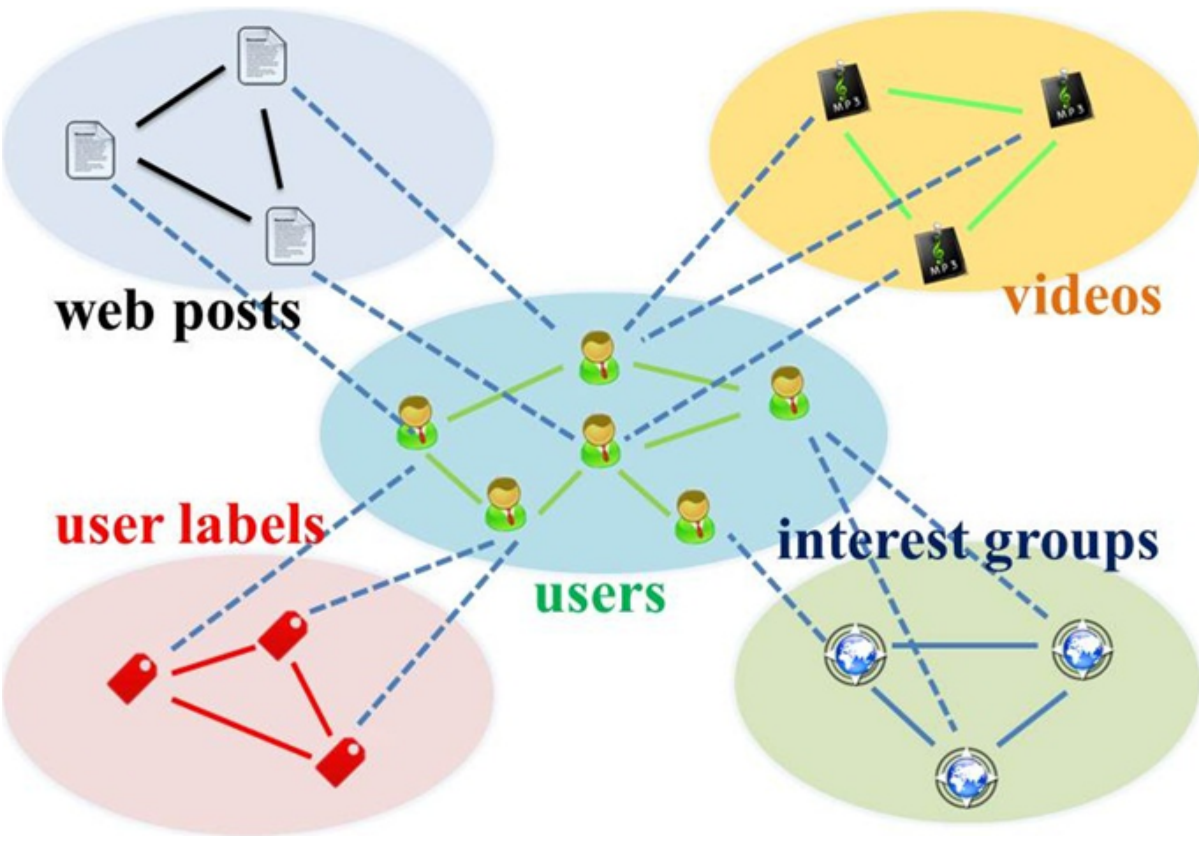
Besides Contexts: Multiple Domains

- Group



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

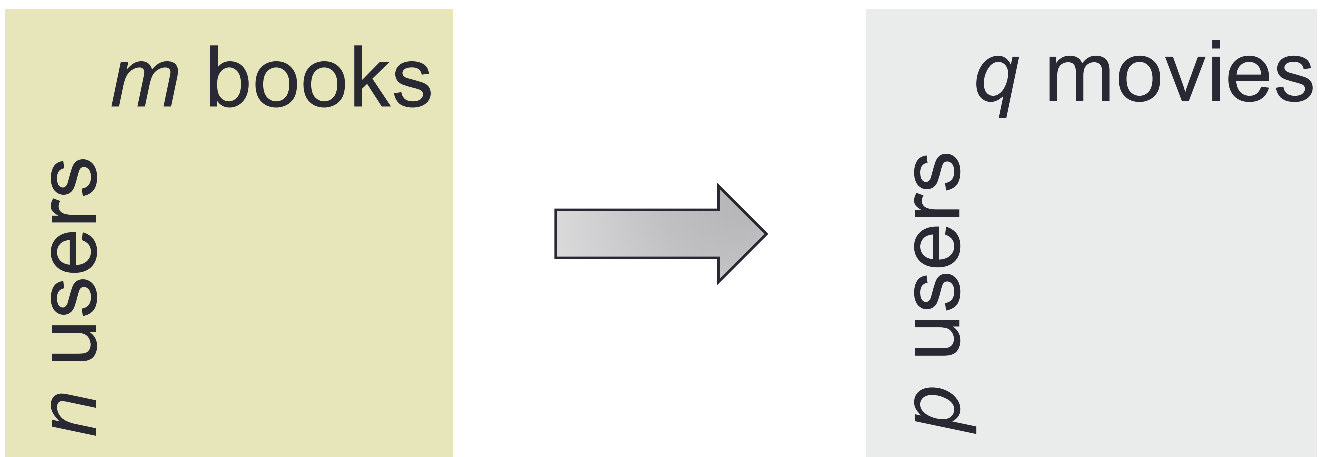
Besides Contexts: Multiple Domains



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

Traditional Cross-Domain CF

- Codebook Transfer (CBT)



$\mathbf{X}_{aux}$ $\mathbf{X}_{tgt}$

Li et al. **Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.** *IJCAI*, 2009.

Traditional Cross-Domain CF

- Codebook Transfer (CBT)

$n \times k$ $\quad m \times l$

$$\min_{\mathbf{U}\geq 0, \mathbf{V}\geq 0, \mathbf{S}\geq 0} \left\| \mathbf{X}_{aux} - \mathbf{U}\mathbf{S}\mathbf{V}^\top \right\|_F^2$$

$$\text{s.t.} \quad \mathbf{U}^\top \mathbf{U} = \mathbf{I}, \mathbf{V}^\top \mathbf{V} = \mathbf{I},$$

m books

n users

$\mathbf{X}_{aux}$

Codebook = User Group × Item Group

$k \times l$

$$\mathbf{B} = [\mathbf{U}_{aux}^\top \mathbf{X}_{aux} \mathbf{V}_{aux}] \oslash [\mathbf{U}_{aux}^\top \mathbf{1}\mathbf{1}^\top \mathbf{V}_{aux}]$$

Li et al. **Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.** *IJCAI*, 2009.

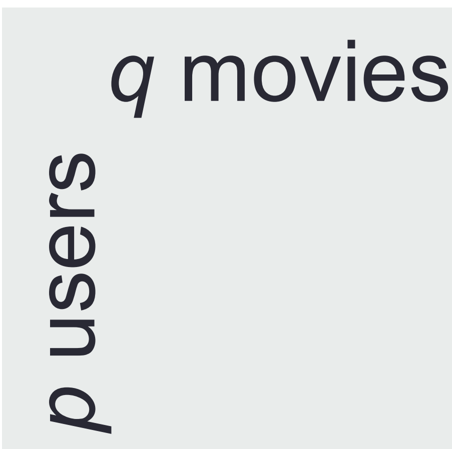
Traditional Cross-Domain CF

- Codebook Transfer (CBT)

Codebook: k × l

p × k

q × l



$\mathbf{X}_{tgt}$

$$\min_{\substack{\mathbf{U}_{tgt}\in\{0,1\}^{p\times k}\\ \mathbf{V}_{tgt}\in\{0,1\}^{q\times l}}} \left\| [\mathbf{X}_{tgt} - \mathbf{U}_{tgt}\mathbf{B}\mathbf{V}_{tgt}^{\top}] \circ \mathbf{W} \right\|_F^2$$

$$\text{s.t.} \quad \mathbf{U}_{tgt}\mathbf{1} = \mathbf{1}, \mathbf{V}_{tgt}\mathbf{1} = \mathbf{1},$$

Li et al. **Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.** *IJCAI*, 2009.

Traditional Cross-Domain CF

❖Codebook Transfer (CBT)

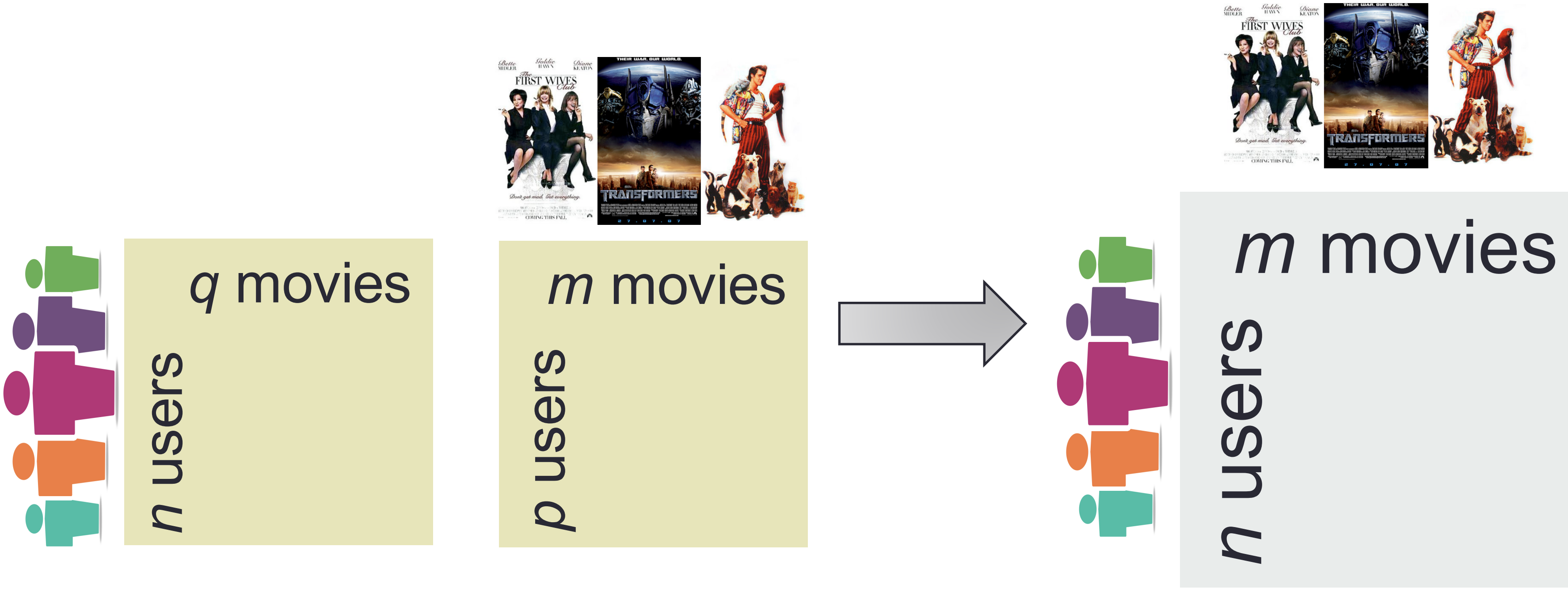
Table 1: MAE on MovieLens (average over 10 splits)

Training Set	Method	Given5	Given10	Given15
ML100	PCC	0.930	0.883	0.873
	CBS	0.874	0.845	0.839
	WLR	0.915	0.875	0.890
	CBT	**0.840**	**0.802**	**0.786**
ML200	PCC	0.905	0.878	0.878
	CBS	0.871	0.833	0.828
	WLR	0.941	0.903	0.883
	CBT	**0.839**	**0.800**	**0.784**
ML300	PCC	0.897	0.882	0.885
	CBS	0.870	0.834	0.819
	WLR	1.018	0.962	0.938
	CBT	**0.840**	**0.801**	**0.785**

Li et al. **Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction.** *IJCAI*, 2009.

Traditional Cross-Domain CF

❖Coordinate System Transfer (CST)



auxiliary data | target data

Pan et al. **Transfer Learning in Collaborative Filtering for Sparsity Reduction.** *AAAI*, 2010.

Traditional Cross-Domain CF

❖Coordinate System Transfer (CST)

Auxiliary data:

$$\min_{\mathbf{U}^{(i)},\mathbf{V}^{(i)},\mathbf{B}^{(i)}} ||\mathbf{Y}^{(i)} \odot (\mathbf{R}^{(i)} - \boxed{\mathbf{U}^{(i)}}\mathbf{B}^{(i)}\boxed{\mathbf{V}^{(i)T}})||_F^2$$

Target data:

$$\min_{\mathbf{U},\mathbf{V},\mathbf{B}} ||\mathbf{Y} \odot (\mathbf{R} - \mathbf{U}\mathbf{B}\mathbf{V}^T)||$$

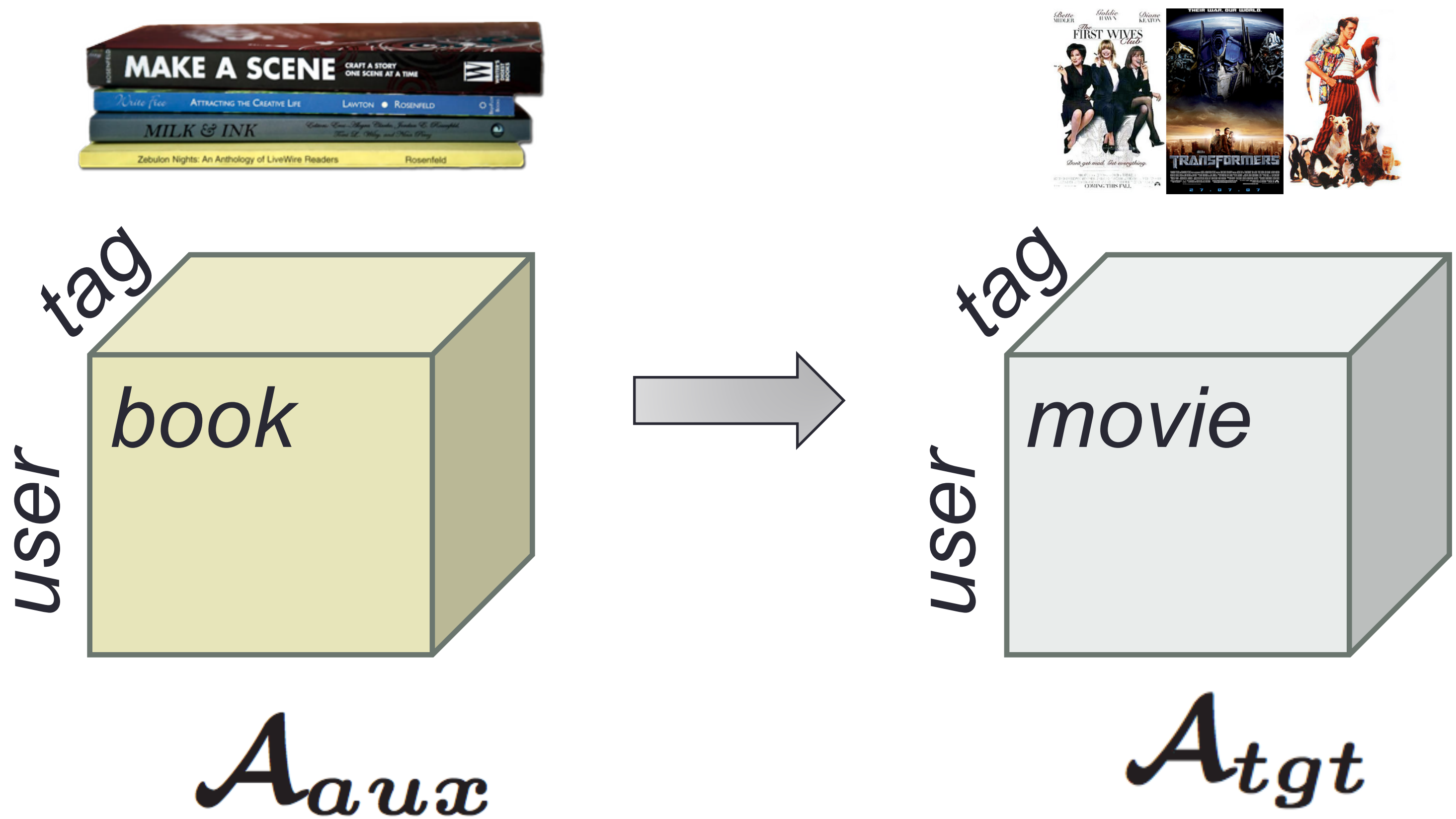
$$+\boxed{\frac{\rho_u}{2}||\mathbf{U} - \mathbf{U}_0||_F^2} + \boxed{\frac{\rho_v}{2}||\mathbf{V} - \mathbf{V}_0||_F^2}$$

$$\text{s.t.} \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}, \mathbf{V}^T\mathbf{V} = \mathbf{I}$$

Pan et al. **Transfer Learning in Collaborative Filtering for Sparsity Reduction.** *AAAI*, 2010.

Traditional Cross-Domain CF

❖FUSE



Traditional Cross-Domain CF

- FUSE

$$\mathcal{A}^{*}_{tgt} = \mathcal{A}^{cluster}_{aux} \times_1 \hat{U}^{(1)}_{tgt} \times_2 \hat{U}^{(2)}_{tgt} \times_3 \hat{U}^{(3)}_{tgt}$$

$$\boldsymbol{f} = \min_{\hat{U}^{(1)}_{tgt} \cdots \hat{U}^{(3)}_{tgt}} \boxed{||\mathcal{A} - \mathcal{A}^{*}_{tgt}||^2_F}$$

$$+\lambda \cdot \sum_{r=1}^{R} tr([\hat{U}^{(1)}_{tgt}]^T (\boldsymbol{D}^{(r)} - \boldsymbol{F}^{(r)}) \hat{U}^{(1)}_{tgt})$$

Traditional Cross-Domain CF

- FUSE

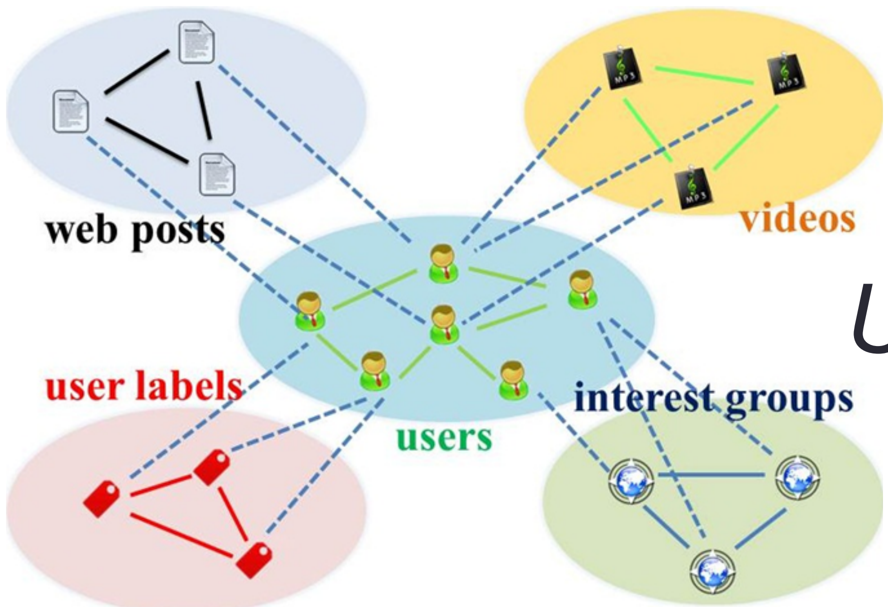
$$[\hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{1})}]_{*r} \leftarrow [\hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{1})}]_{*r} \circledast \frac{[\boldsymbol{A}_{(1)} \boldsymbol{S}_{(1)}^{T}]_{*r} + \lambda \boldsymbol{F}^{(r)} [\hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{1})}]_{*r}}{[\hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{1})} \boldsymbol{S}_{(1)} \boldsymbol{S}_{(1)}^{T}]_{*r} + \lambda \boldsymbol{D}^{(r)} [\hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{1})}]_{*r}}$$

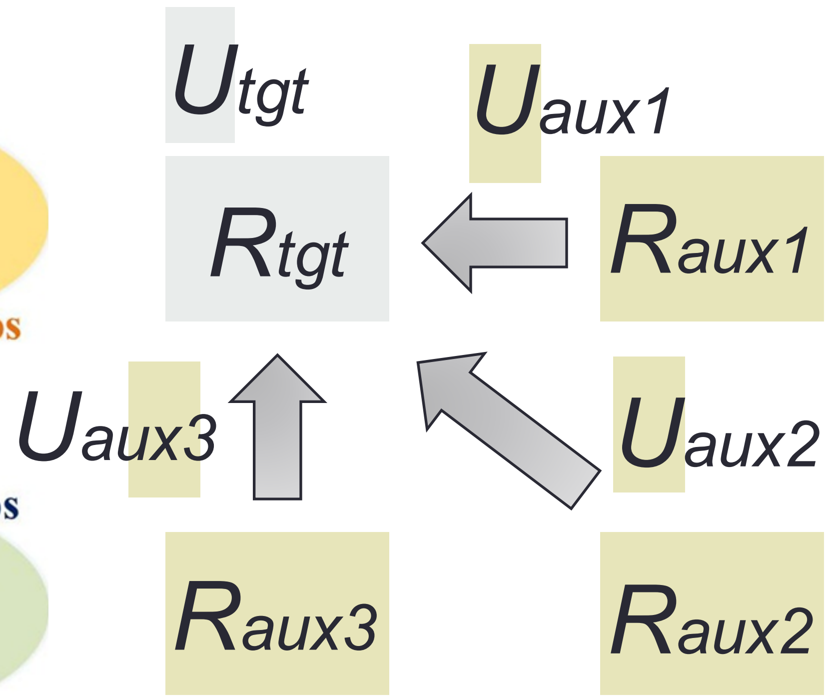
Gradient Descent Methods

$$\hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{2})} \leftarrow \hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{2})} \circledast \frac{\boldsymbol{A}_{(2)} \boldsymbol{S}_{(2)}^{T}}{\hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{2})} \boldsymbol{S}_{(2)} \boldsymbol{S}_{(2)}^{T}}$$

$$\hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{3})} \leftarrow \hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{3})} \circledast \frac{\boldsymbol{A}_{(3)} \boldsymbol{S}_{(3)}^{T}}{\hat{\boldsymbol{U}}_{\boldsymbol{tgt}}^{(\boldsymbol{3})} \boldsymbol{S}_{(3)} \boldsymbol{S}_{(3)}^{T}}$$

When Social Recommendation Meets Multiple Domains

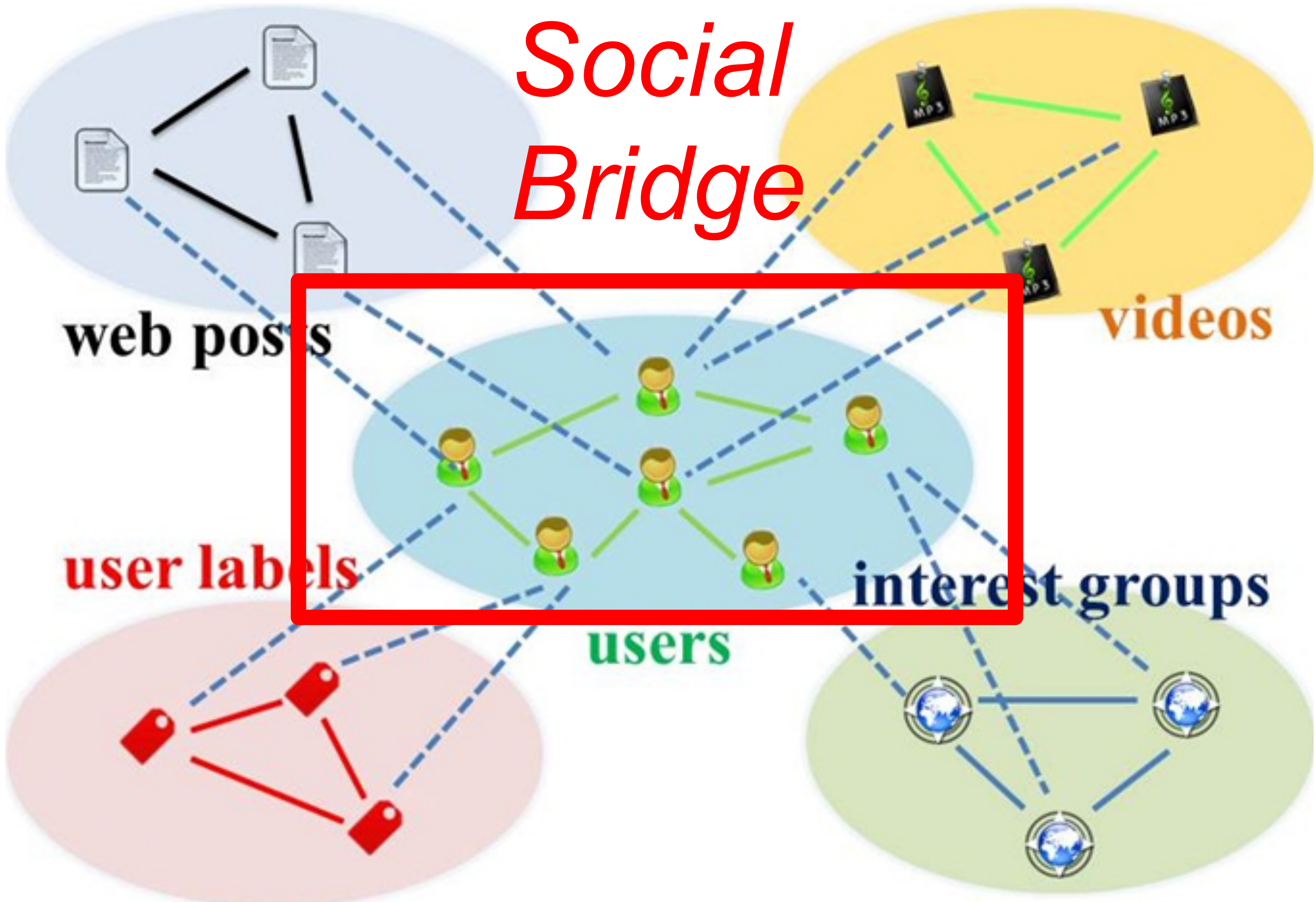




Bridge: User × Cluster

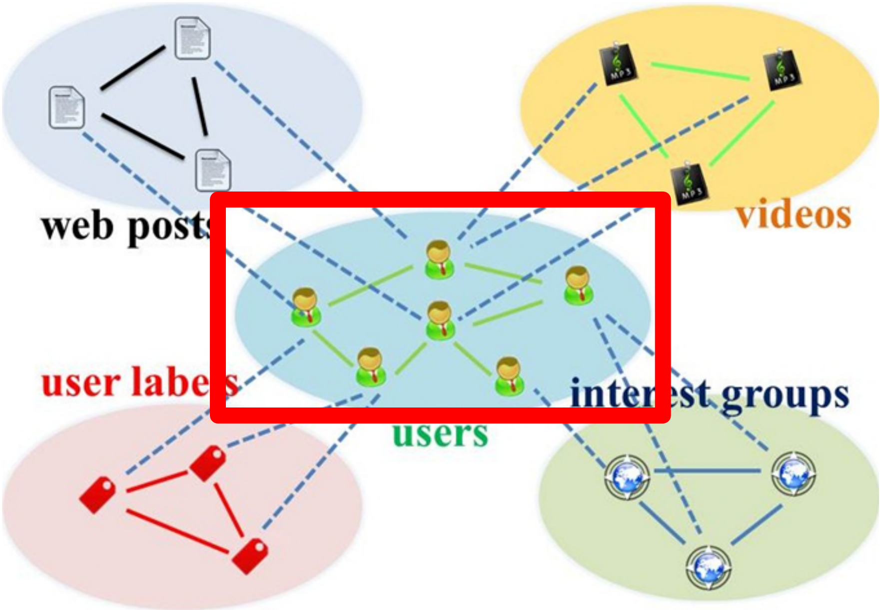
Jiang et al. **Social Recommendation across Multiple Relational Domains.** *CIKM*, 2012. **Social Recommendation with Cross-Domain Transferable Knowledge.** *TKDE*, 2015.

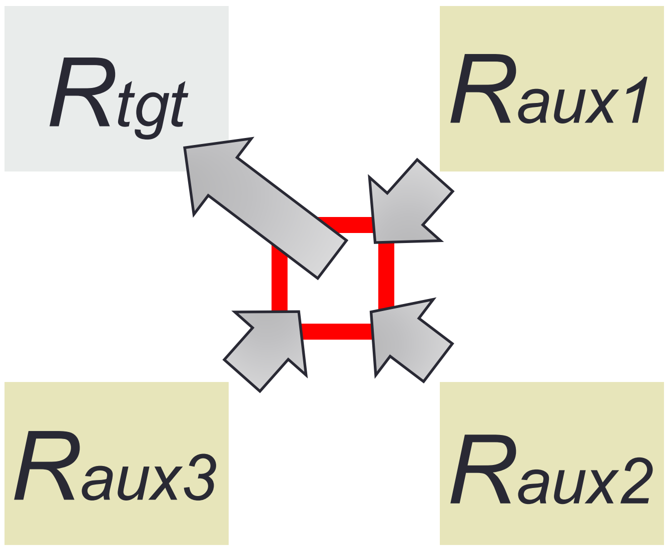
When Social Recommendation Meets Multiple Domains



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

When Social Recommendation Meets Multiple Domains



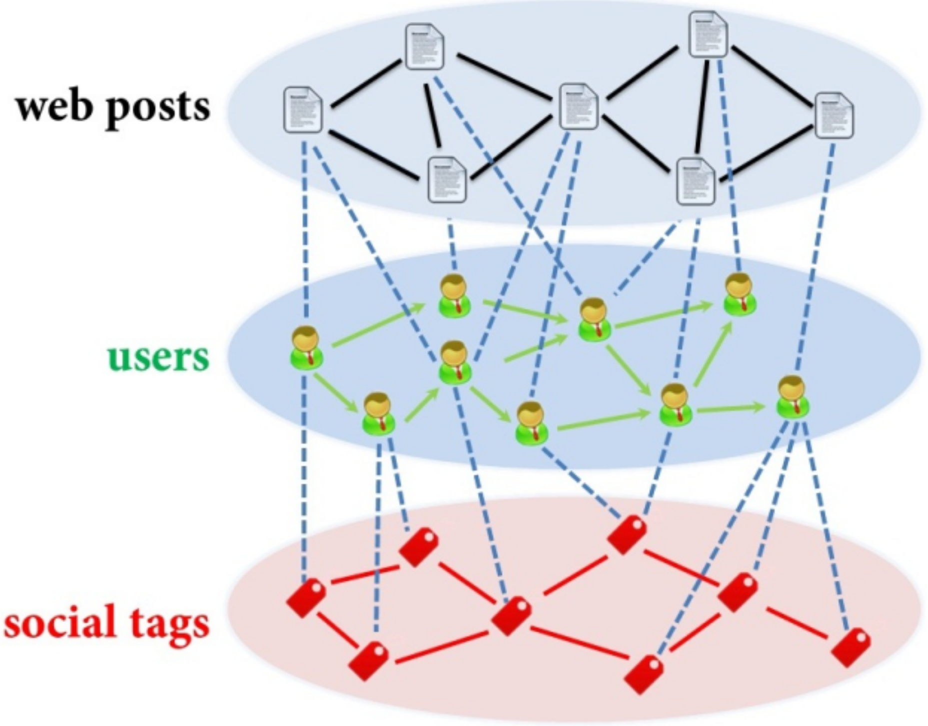


Bridge: User × User

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

Hybrid Random Walk

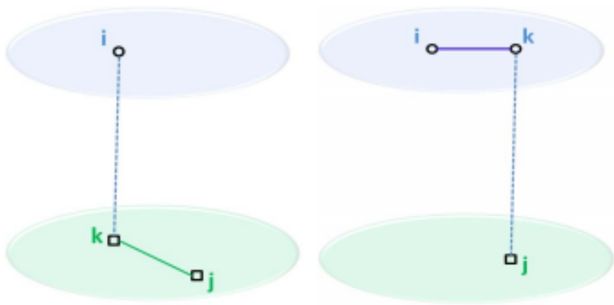
❖Starting with a Second-Order Start-Structured Graph



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

Hybrid Random Walk

❖Updating cross-domain links



(a) Route 1 (b) Route 2

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

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

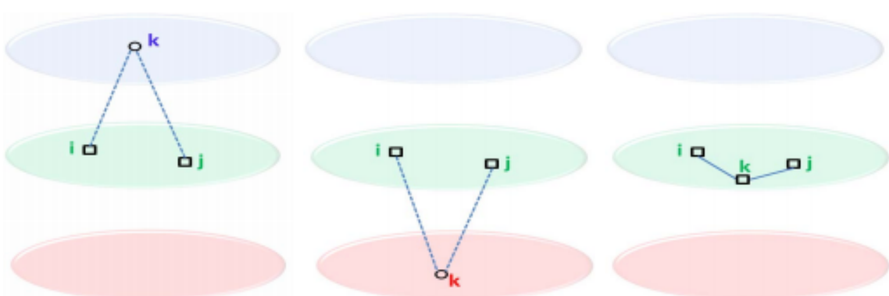
$$p_{ij}^{(\mathcal{UT})^+} = \eta \sum_{u_k \in \mathcal{U}} r_{ik}^{(\mathcal{U})} p_{kj}^{(\mathcal{UT})^+} + (1-\eta) \sum_{t_k \in \mathcal{T}} p_{ik}^{(\mathcal{UT})^+} r_{kj}^{(\mathcal{T})}$$

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

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

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

❖Updating within-domain links



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

$$r_{ij}^{(\mathcal{U})} = \tau^{(\mathcal{P})} \Big(\mu \sum_{p_k \in \mathcal{P}} p_{ik}^{(\mathcal{UP})^+} p_{jk}^{(\mathcal{UP})^+} + (1-\mu) \sum_{p_k \in \mathcal{P}} p_{ik}^{(\mathcal{UP})^-} p_{jk}^{(\mathcal{UP})^-}\Big)$$
$$+ \tau^{(\mathcal{T})} \sum_{t_k \in \mathcal{T}} p_{ik}^{(\mathcal{UT})^+} p_{jk}^{(\mathcal{UT})^+} + \tau^{(\mathcal{U})} \sum_{u_k \in \mathcal{U}} r_{ik}^{(\mathcal{U})} r_{kj}^{(\mathcal{U})} \qquad (12)$$

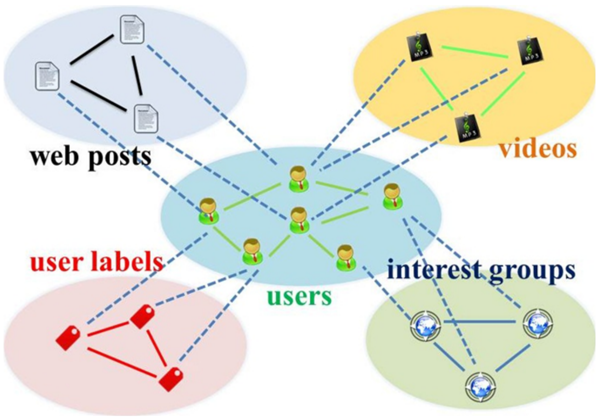
$$\mathbf{R}^{(\mathcal{U})}(t+1) =$$
$$\tau^{(\mathcal{P})} \Big(\mu \mathbf{P}^{(\mathcal{UP})^+}(t) \mathbf{P}^{(\mathcal{UP})^+}(t)^T + (1-\mu) \mathbf{P}^{(\mathcal{UP})^-}(t) \mathbf{P}^{(\mathcal{UP})^-}(t)^T\Big)$$
$$+ \tau^{(\mathcal{T})} \mathbf{P}^{(\mathcal{UT})^+}(t) \mathbf{P}^{(\mathcal{UT})^+}(t)^T + \tau^{(\mathcal{U})} \mathbf{R}^{(\mathcal{U})}(t) \mathbf{R}^{(\mathcal{U})}(t)^T \qquad (13)$$

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

Hybrid Random Walk

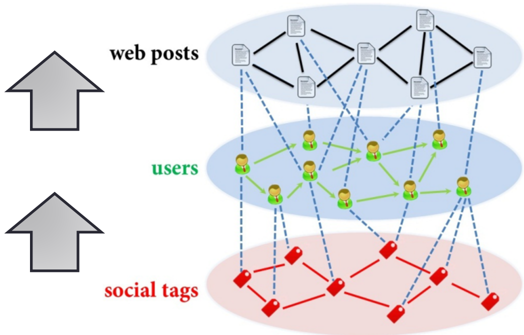
❖High-Order Star-Structured Graph

$$
\begin{aligned}
\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t+1) &= \delta_i \mathbf{R}^{(\mathcal{U})}(t)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t) + (1-\delta_i)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t)\mathbf{R}^{(\mathcal{D}_i)} \\
\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t+1) &= \delta_i \mathbf{R}^{(\mathcal{U})}(t)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t) + (1-\delta_i)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t)\mathbf{R}^{(\mathcal{D}_i)} \\
\mathbf{R}^{(\mathcal{U})}(t+1) &= \sum_{\mathcal{D}_i \in \mathcal{D}} \tau_i \mu_i \mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^+}(t)^T \\
&+ \sum_{\mathcal{D}_i \in \mathcal{D}} \tau_i (1-\mu_i)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t)\mathbf{P}^{(\mathcal{U}\mathcal{D}_i)^-}(t)^T \\
&+ \tau^{(\mathcal{U})}\mathbf{R}^{(\mathcal{U})}(t)\mathbf{R}^{(\mathcal{U})}(t)^T \qquad (20)
\end{aligned}
$$



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

Hybrid Random Walk



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

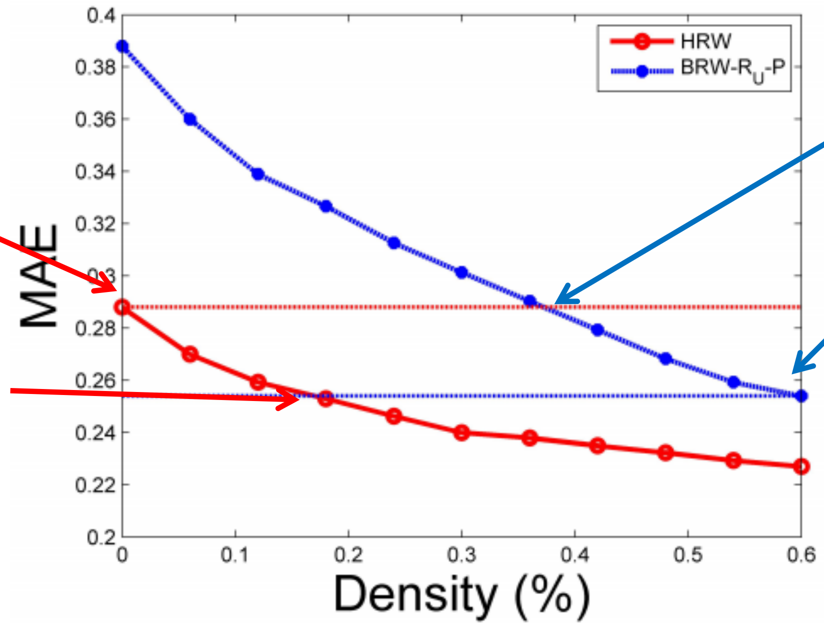
Jiang et al. **Social Recommendation across Multiple Relational Domains.** *CIKM*, 2012. **Social Recommendation with Cross-Domain Transferable Knowledge.** *TKDE*, 2015.

Hybrid Random Walk

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

0 user-post
100% user-label

18% user-post
100% user-label



35% user-post
0 user-label

60% user-post
0 user-label

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

Besides Cross-Domain…

- How about Cross-Platform?
- Partially aligned/overlapped users!

Little is Much: Bridging Cross-Platform Behaviors through Overlapped Crowds
by M. Jiang, P. Cui, N. J. Yuan, X. Xie, S. Yang. AAAI Conference on Artificial Intelligence (AAAI) , 2016.

Questions for Modeling Individual Behavior

- What is individual behavior in social networks?
- Why should we study individual behavior?
- What are the state-of-the-art models?
  - Modeling behaviors and social relations
  - Modeling social contexts
  - Modeling spatiotemporal contexts
  - Modeling multiple domains in social networks

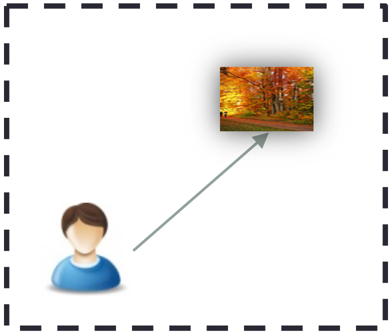
Summary for Modeling Individual Behavior

- Like, Reply, Share, Retweet, Favorite, Comment …
- Pattern discovery, prediction and social recommendation
- Memory based social recommenders
  - TidalTrust, MoleTrust, TrustWalker
- Model based social recommenders
  - SoRec, "Social Trust" Ensemble, SoReg
- ContextMF: Social contexts (preference & influence)
- FEMA: Spatiotemporal contexts (multi-faceted & dynamic)
- Traditional cross-domain CF
  - CBT, CST, FUSE
- Hybrid Random Walk: Social bridging multiple domains

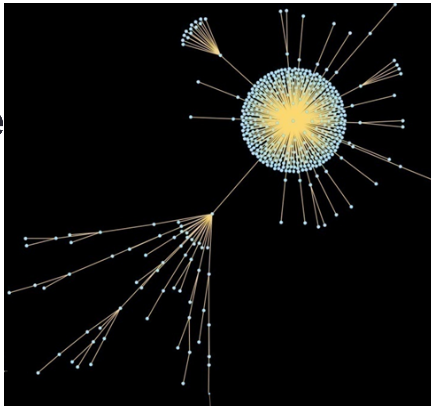
Outline

- **Prediction for natural behavior**
  - Modeling individual behavior (MICRO)
  - **Modeling information cascade (MACRO)**
- Detection for unnatural behavior
  - Suspicious behavior detection

From Micro to Macro



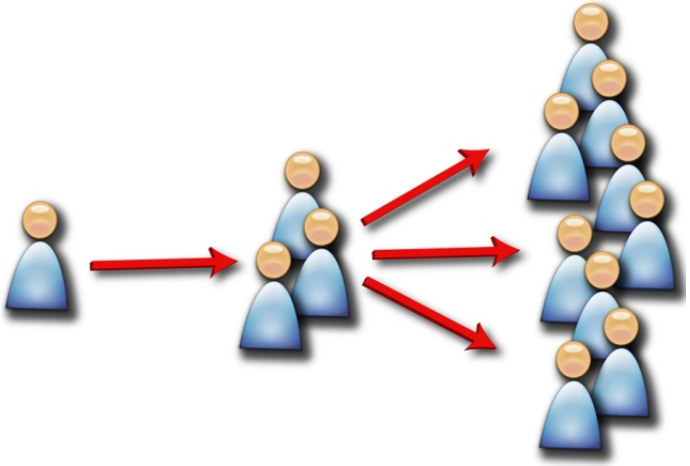
Dynamically Aggregate



Information spreading is a macro phenomenon which is driven by individual user behaviors in microscopic level.

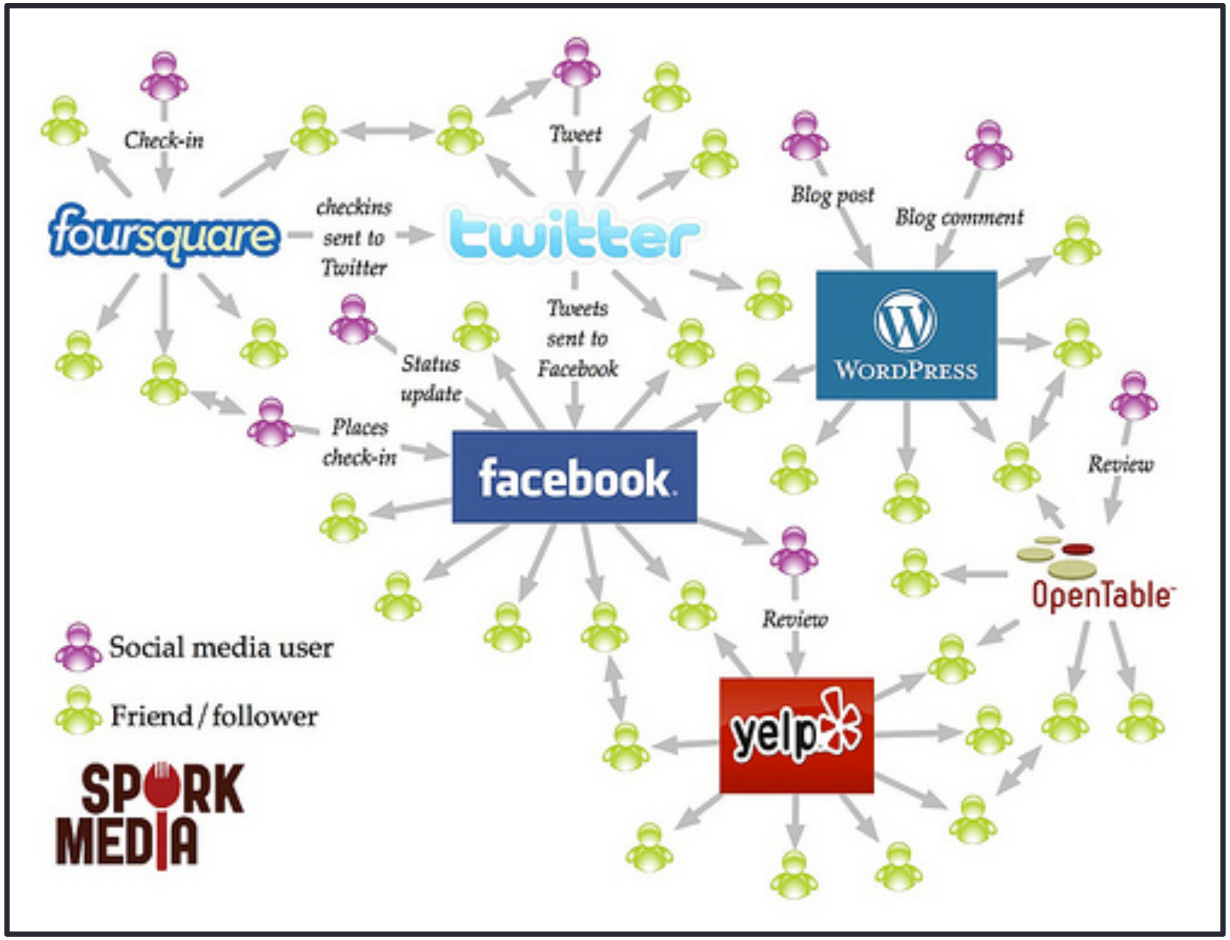
Cascades: Information Spreading

- In network environment, if decentralized nodes act on the basis of how their neighbors act at earlier time, ***cascades*** will be formed.
  - Word-of-mouth
  - Cascading
  - Diffusion
  - Propagation



Information Spreading is Ubiquitous

Social Media



Information spreading is the major way of communication in social media.

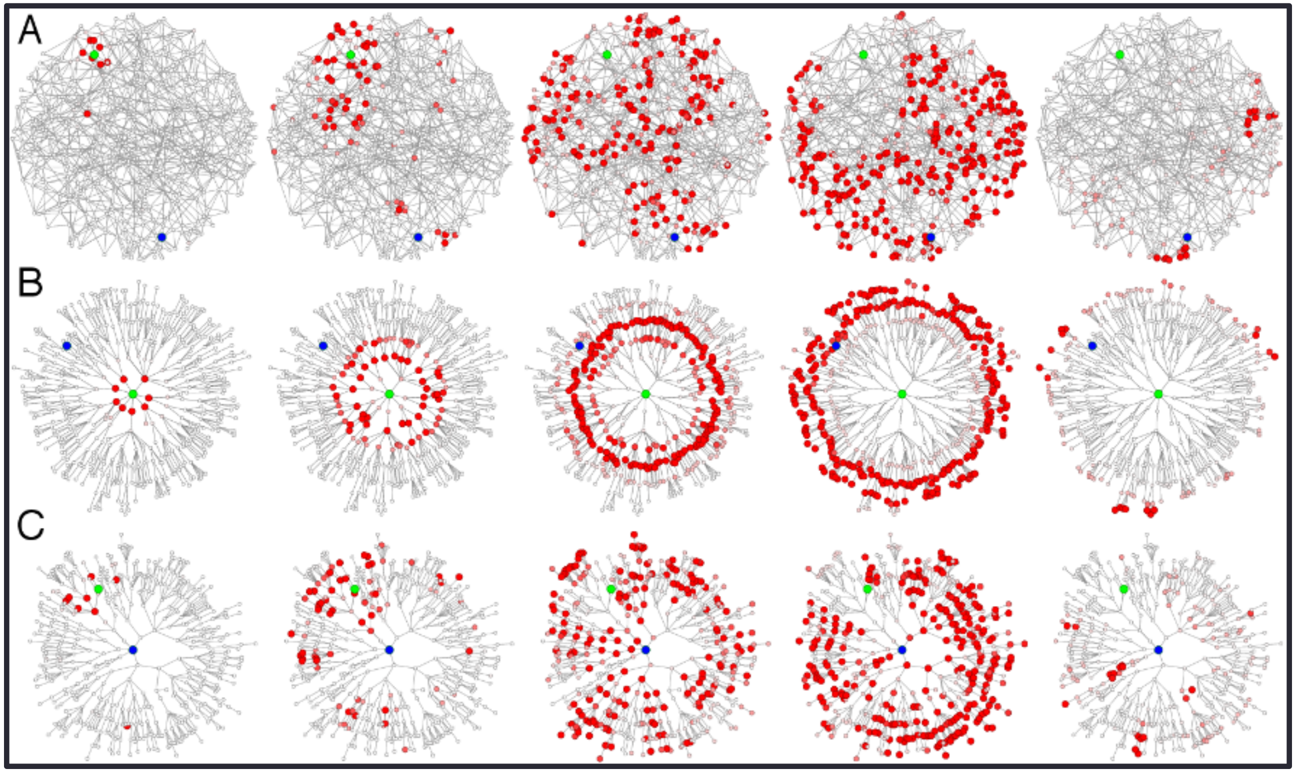
Information Spreading is Ubiquitous

Word-of-Mouth (Marketing)

How to fully exploit the power of word-of-mouth in marketing?

Information Spreading is Ubiquitous

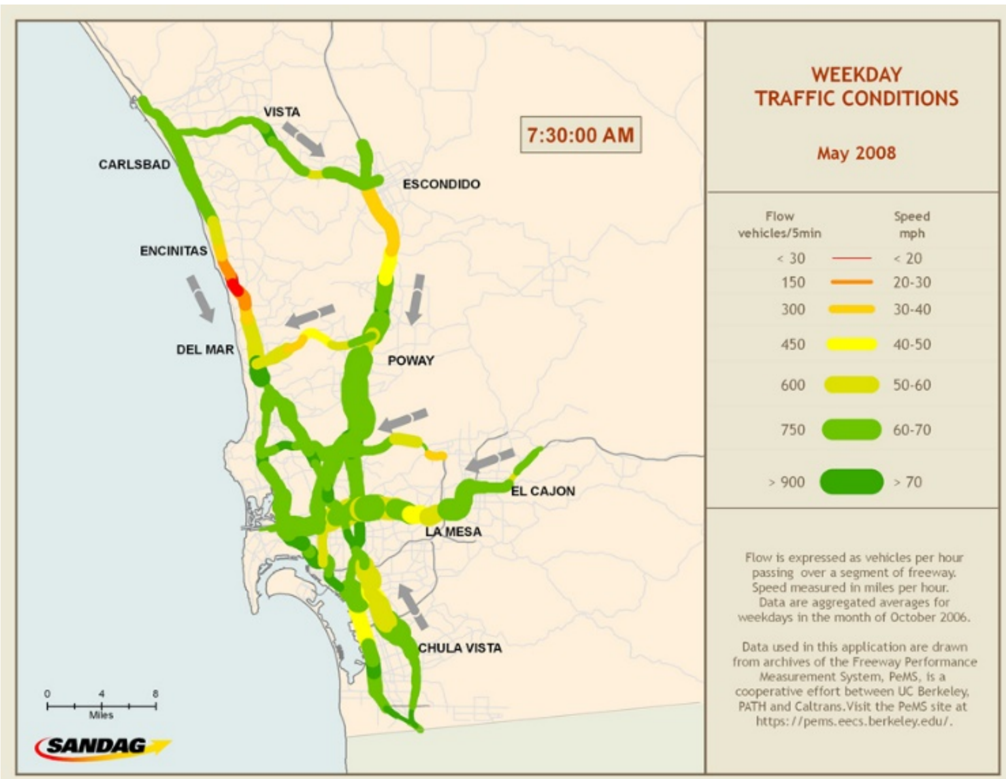
Epidemics



Share similar dynamic process as information spreading.

Information Spreading is Ubiquitous

Traffic



Traffic jams spread through road network. How to model, predict and intervene?

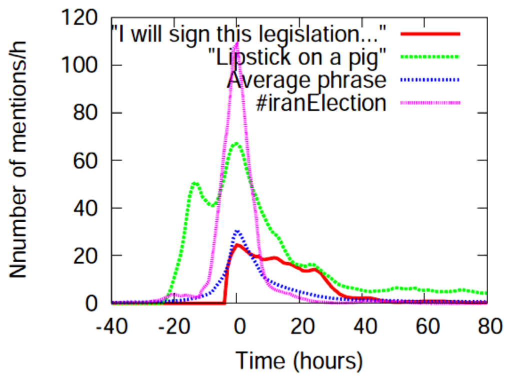
How to understand information spreading mechanism, and furthermore, predict the information spreading process?

Related Research

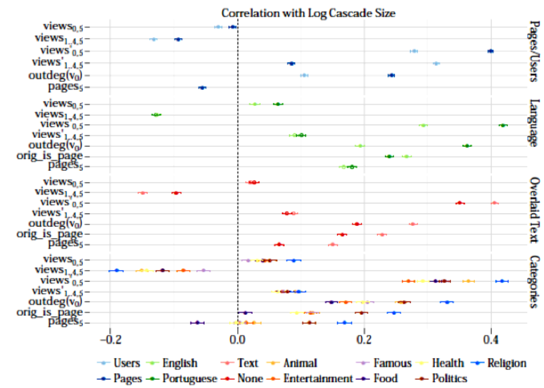
Understand

Prediction

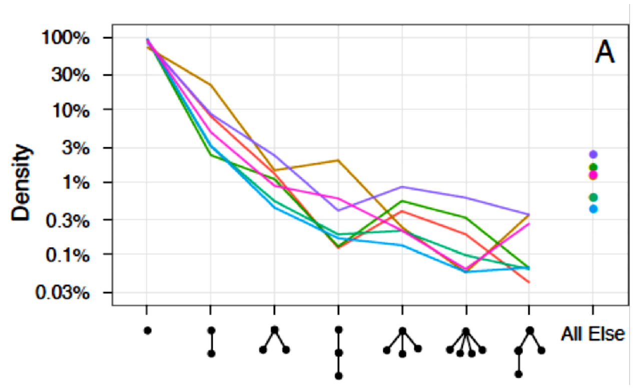
Rise and fall patterns in cascading curve



Popularity prediction



Structural units of information cascades



Regard cascades as a whole and extract cascade-level features for understanding and prediction.

Macro Phenomenon v.s. Micro Mechanism

Information spreading is driven by a cascade of user adoption behaviors.

Micro		Macro
Behavioral Dynamics	→	Information Spreading

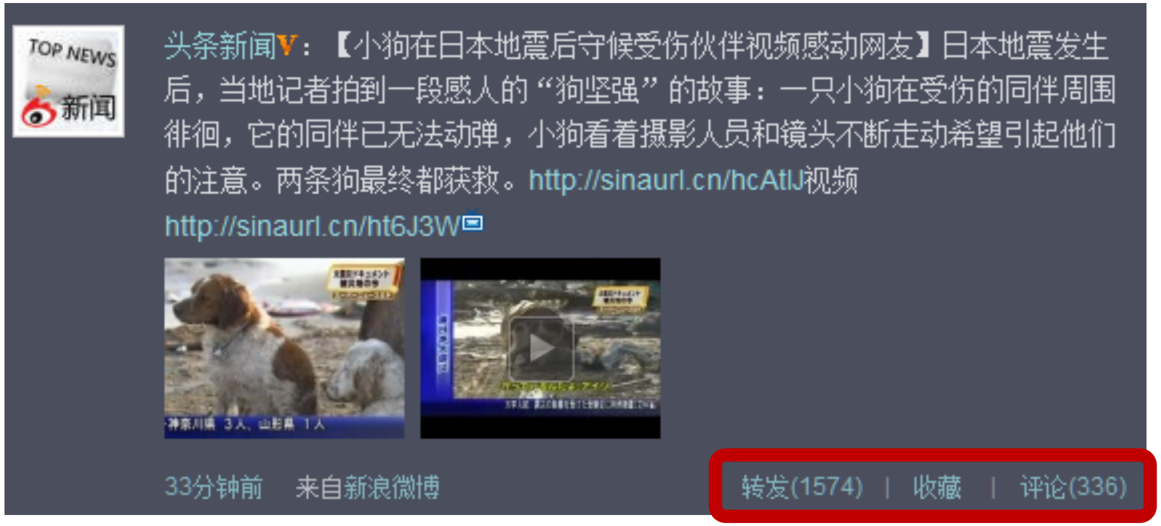
Behavior-Driven Information Spreading Modeling

Ultimate Goal: Bridge the gap between macro phenomena of information spreading and micro behavioral mechanism.

One-Hop Cascade Prediction	Cascading Outbreak Prediction	Dynamic Process Prediction
Predict the collective response of a user's followers	Predict whether the information will break out in future	Predict the dynamic cascading process of a piece of information
SIGIR'11, AAAI'11	KDD'13	ICDM'15

The problem:

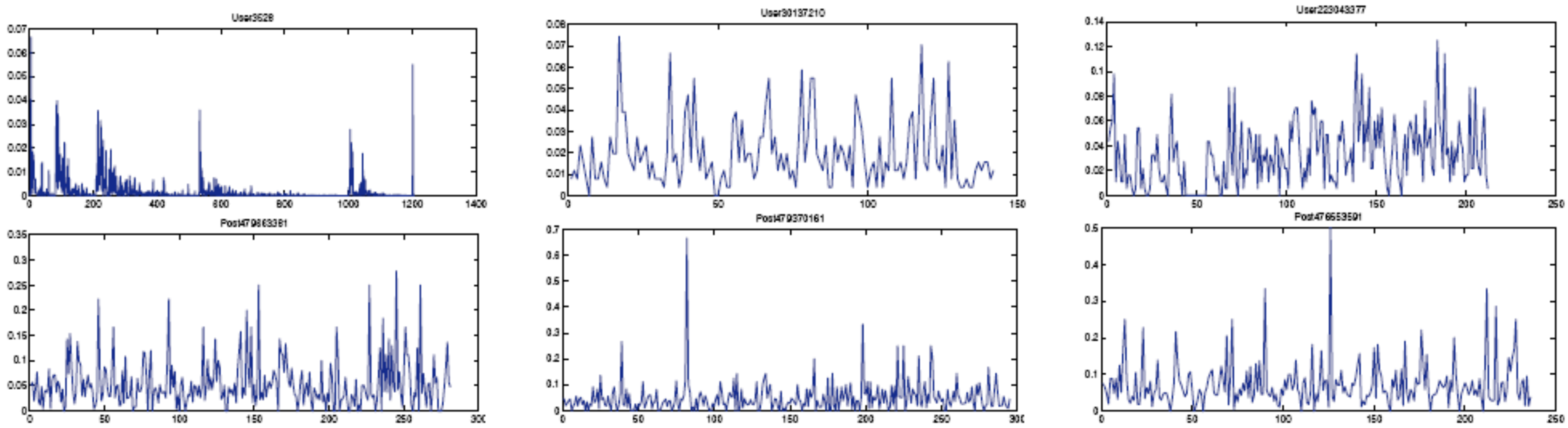
To predict the percentage of a user’s followers that will retweet the microblog after the user retweet it.



The Dimensions

Are big users always trigger high forwarding numbers?

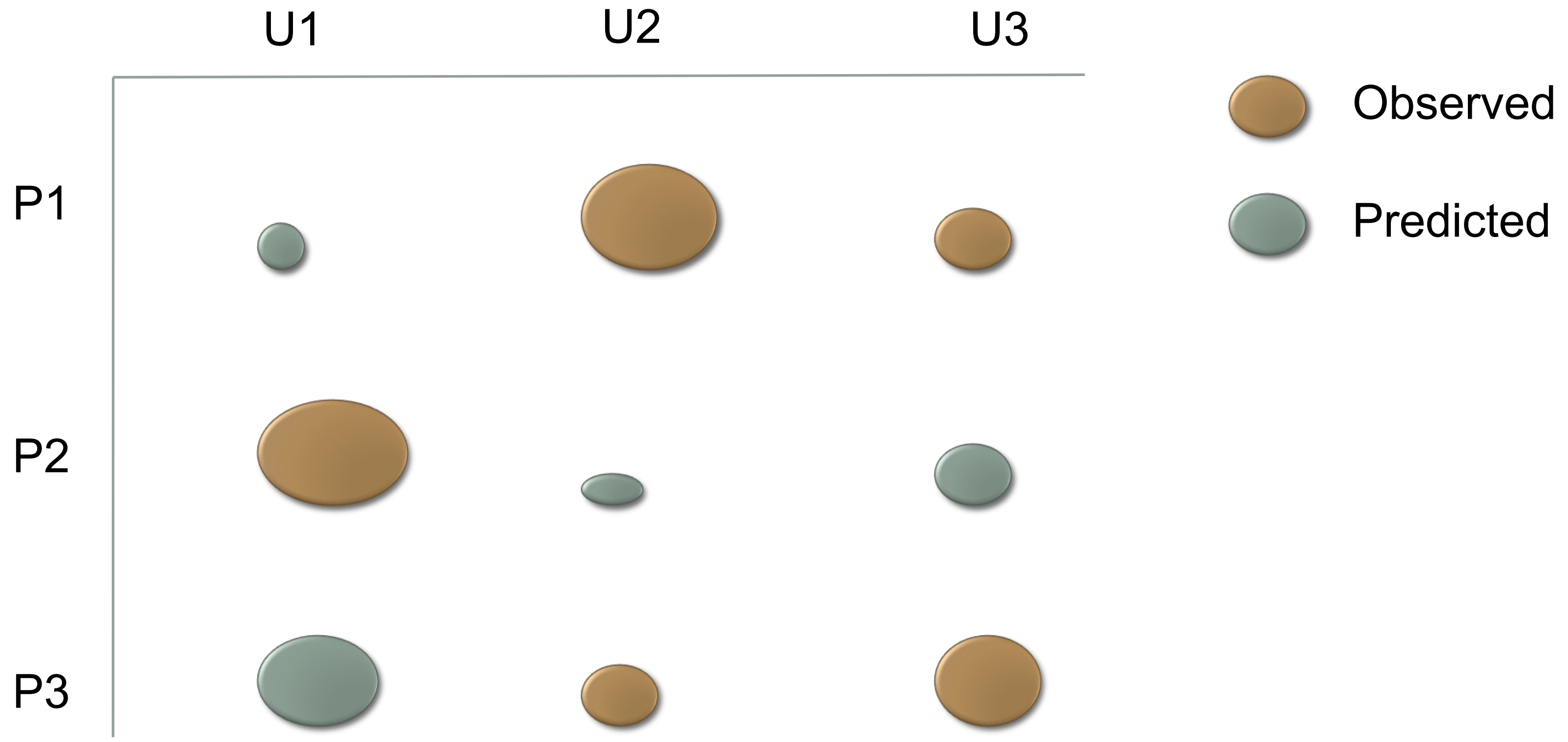
Post Variance



User Variance

Are popular tweets always trigger high forwarding numbers?

Problem Formulation



✓Given an user, rank the web posts to share
✓Given a web post, rank the users to target

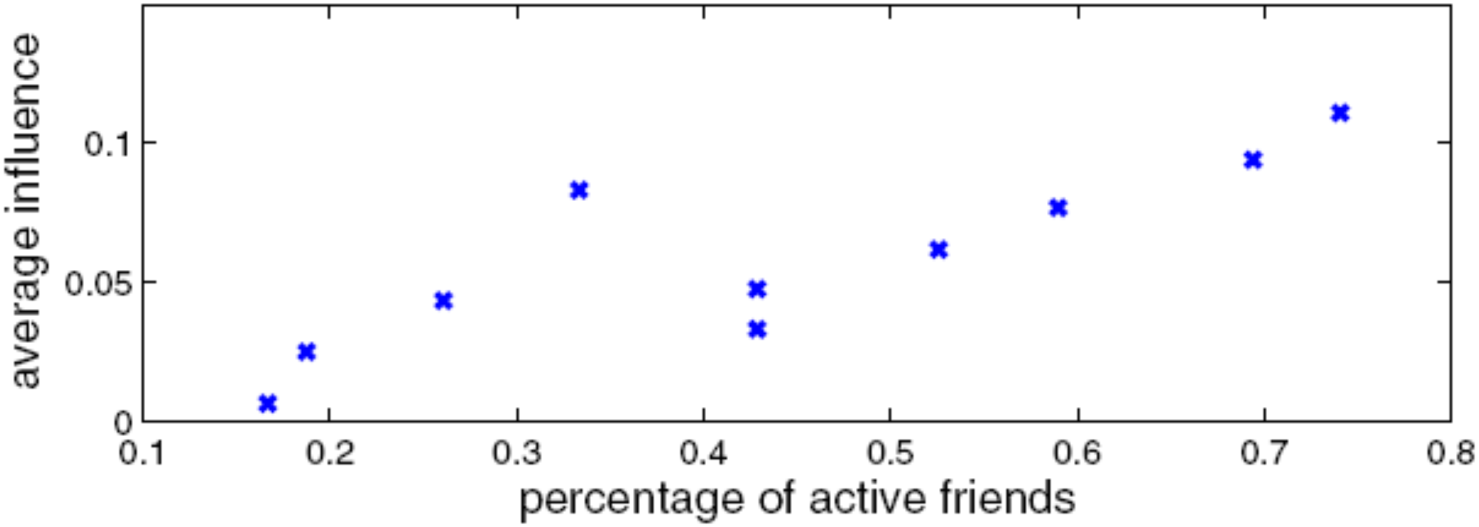
Density 0.1%

We need priors on users and posts.

Predictive Factors

Percentage of active friends

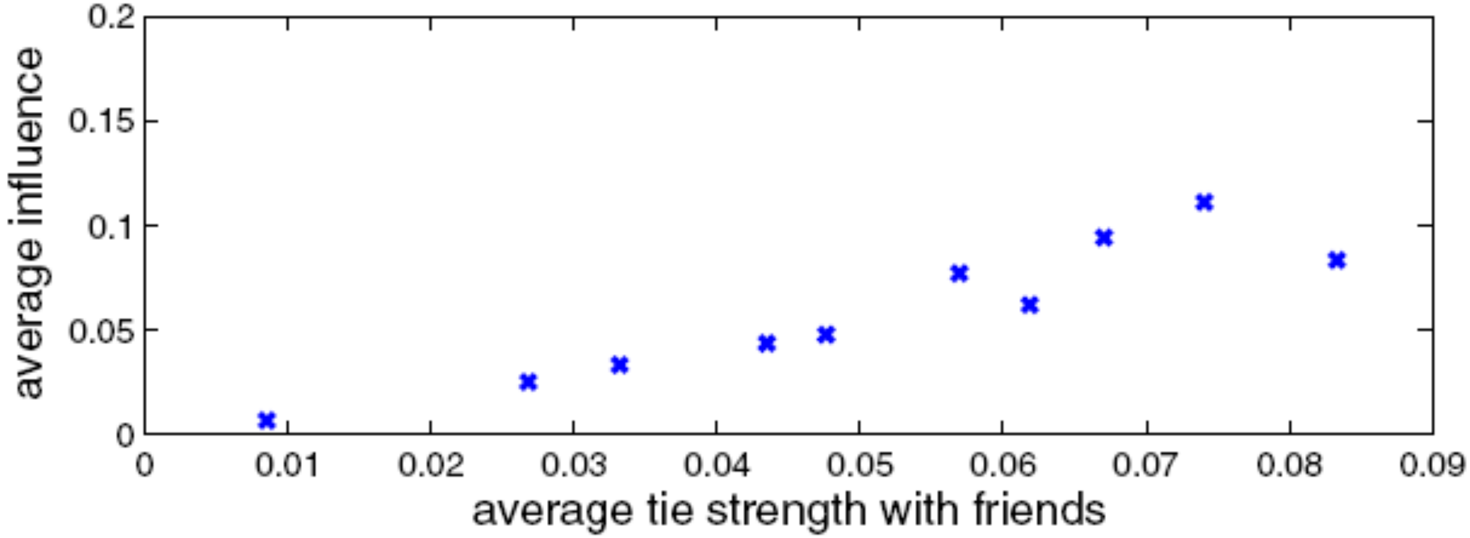
$$uf_1(u_i) = \frac{\sum_{u_r \in \mathcal{N}(u_i)} \delta(act(u_r) \geq \tau)}{|\mathcal{N}(u_i)|}$$



Predictive Factors

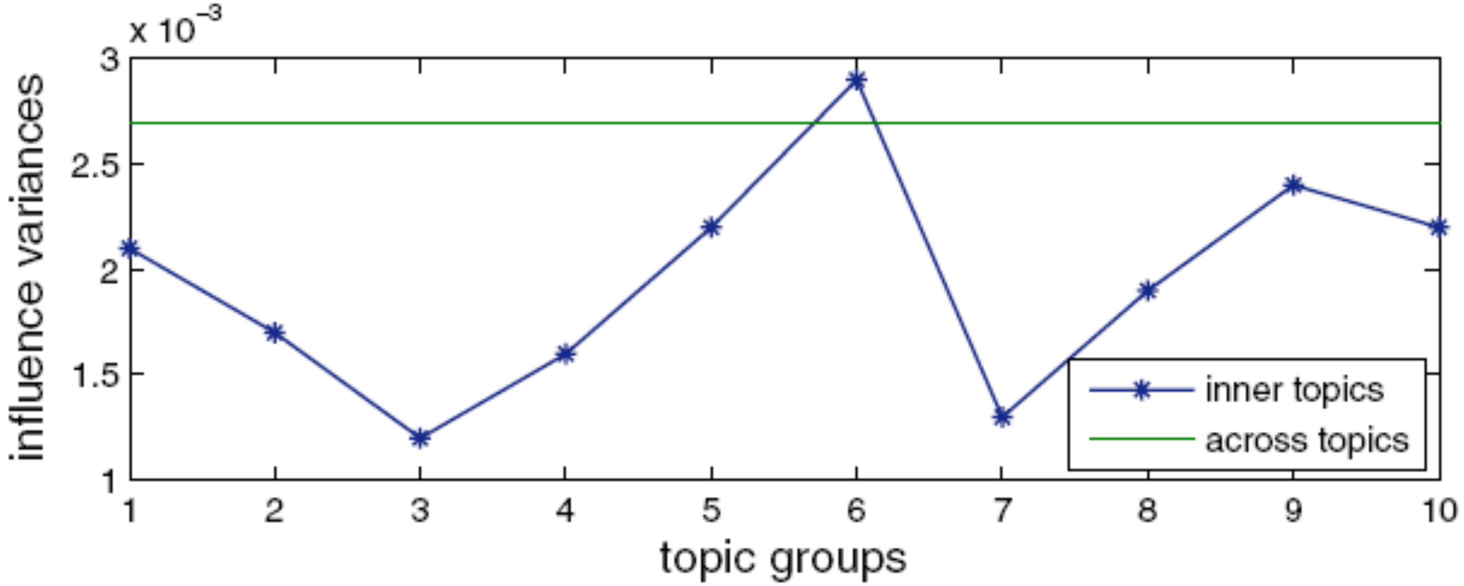
Average tie strength

$$uf_2(u_i) = \frac{\sum_{u_r \in \mathcal{N}(u_i)} \frac{tie(u_i, u_r)}{\sum_j Y_{ij}}}{|\mathcal{N}(u_i)|}$$



Predictive Factors

The introduction of post topic groups can reduce the variances of influences.



Modeling

Baseline objective function

$$\min_{\mathbf{U},\mathbf{V}} \left\|\mathbf{X} - \mathbf{U}\mathbf{V}^\top\right\|_F^2 + \gamma \|\mathbf{U}\|_F^2 + \delta \|\mathbf{V}\|_F^2$$
$$s.t. \mathbf{U} \geq 0, \ \mathbf{V} \geq 0$$

We suppose the users with similar observed predictive factors have similar distribution in latent space

$$\mathcal{J}_3 = \left\|\underbrace{\mathbf{W}}_{\text{User similarity matrix}} - \mathbf{U}\mathbf{U}^\top\right\|_F^2$$

We constrain the latent post space by topic distributions

$$\mathcal{J}_4 = \left\|\underbrace{\mathbf{C}}_{\text{Post content matrix}} - \mathbf{V}\underbrace{\mathbf{G}}_{\text{Topic matrix}}^\top\right\|_F^2$$

Modeling

Hybrid Factor Non-Negative Matrix Factorization (HF-NMF)

$$\begin{aligned}
\min_{\mathbf{U},\mathbf{V},\mathbf{G}} \quad & \left\|\mathbf{X}-\mathbf{U}\mathbf{V}^{\top}\right\|_F^2 + \alpha\left\|\mathbf{W}-\mathbf{U}\mathbf{U}^{\top}\right\|_F^2 + \beta\left\|\mathbf{C}-\mathbf{V}\mathbf{G}^{\top}\right\|_F^2 \\
& +\gamma\left\|\mathbf{U}\right\|_F^2 + \delta\left\|\mathbf{V}\right\|_F^2 \\
s.t. \quad & \mathbf{U}\geq 0,\ \mathbf{V}\geq 0,\ \mathbf{G}\geq 0 \qquad (12)
\end{aligned}$$

Ranking Criterion

	User	Ranking	Post	Ranking
	η	ϱ	η	ϱ
HF-NMF	**0.8942**	**0.9389**	**0.8012**	**0.8697**
bNMF+UF	0.8739	0.9088	0.7423	0.8334
bNMF+PF	0.8236	0.8412	0.7654	0.8548
bNMF	0.813	0.8342	0.7358	0.7926
AvgU	0.7824	0.8056	0.7047	0.7583
AvgP	0.6973	0.7143	0.6746	0.736
CoxPH	0.6596	0.6893	0.659	0.6762
LR	0.6524	0.697	0.6328	0.6593

The advantages of HF-NMF is more apparent in ranking evaluations.

Examples

For a user, ranking the posts

PostIDs	8783	9993	6551	8169	3550	8698	1404	5655	7825	4459
RankOrder(groundtruth)	1	2	3	4	5	6	7	8	9	10
SocialInfluence(groundtruth)	73	53	53	33	13	13	13	13	6	6
RankOrder(Prediction)	1	3	2	4	9	6	7	8	5	10
SocialInfluence(Prediction)	65	43	44	31	12	20	15	14	25	9

For a post, ranking the users

UserIDs	2627	1287	2336	2952	4466	2764	3052	0893	7666	4909
RankOrder(groundtruth)	1	2	3	4	5	6	7	8	9	10
SocialInfluence(groundtruth)	33	26	19	19	13	13	6	6	6	6
RankOrder(Prediction)	4	1	2	3	5	6	7	8	9	10
SocialInfluence(Prediction)	16	27	19	17	13	11	7	6	6	6

Discussions

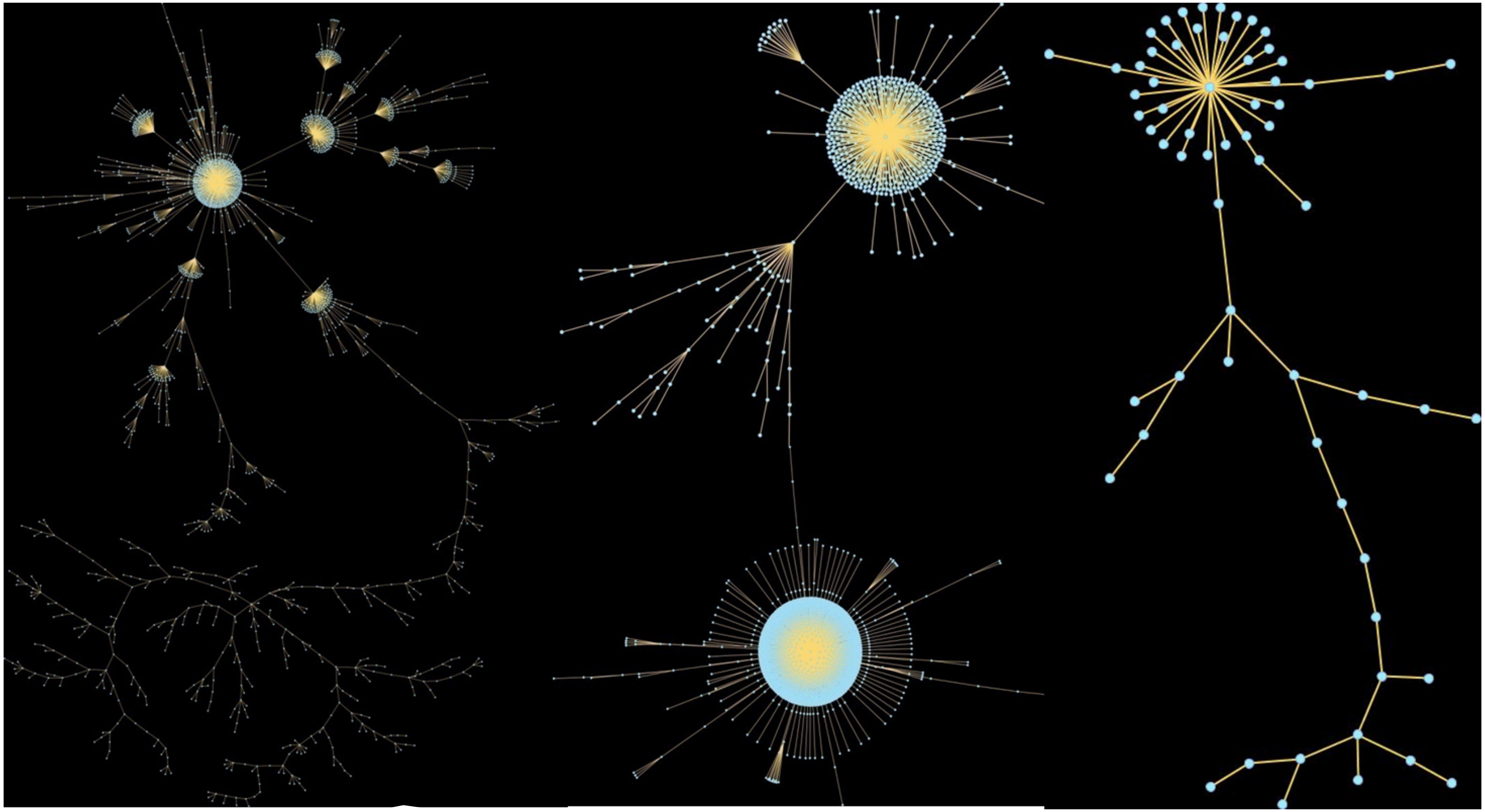
- The collective retweeting behaviors of a user's followers is predictable in fine granularity.
- Can we use the results of one-hop cascade prediction to predict the whole cascades? No!
    - Inapplicable in real applications
    - Error aggregation
- **Hint:** Different users play different roles in information spreading.

Predictive Modeling on Information Spreading

Ultimate Goal: Bridge the gap between macro phenomena of information spreading and micro behavioral mechanism.

One-Hop Cascade Prediction	Cascading Outbreak Prediction	Dynamic Process Prediction
Predict the collective response of a user's followers	Predict whether the information will break out in future	Predict the dynamic cascading process of a piece of information
SIGIR'11, AAAI'11	KDD'13	ICDM'15

Cascading Outbreak Prediction



Can we **predict** whether a tweet will be hot in future?

Outbreak prediction

- Basic Hypothesis: User behaviors cause outbreaks
- Experience: Different users play different roles in causing outbreaks
- How to identify the important users?
  - Topology measures
    - Indegree, centralities, etc.
  - Influential nodes
    - Suppose the cascading process

But does the real data follow the hypothesized cascading process and topology measures?

A Data Driven Approach

- Mining from massive historical data

Early stage users as features

Paths as samples

$$\mathrm{M}\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \rightarrow Result \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

Outbreak signals

Selected as sensors
to predict outbreaks

Challenges

- The outbreak prediction and node selection procedures need to be jointly optimized
- The node selection need to be parsimonious so that the monitoring over the selected sensors can be cost effective
- The node selection process need to be efficient so that the method can be applied into large realistic networks

Orthogonal Sparse LOgistic Regression (OSLOR)

$$L(\boldsymbol{\theta}) = h(\mathbf{X}_{i\cdot})^{y_i} \cdot (1 - h(\mathbf{X}_{i\cdot}))^{1-y_i}$$

$$\log L(\boldsymbol{\theta}) = -\sum_{i=1}^{m}(\log(1 + e^{\mathbf{X}_{i\cdot}\boldsymbol{\theta}})) + \mathbf{y}^{\top}\mathbf{X}\boldsymbol{\theta}$$

$$F(\boldsymbol{\theta}) = T_1(\boldsymbol{\theta}) + T_2(\boldsymbol{\theta}) + T_3(\boldsymbol{\theta})$$

$$T_1(\boldsymbol{\theta}) = -\log L(\boldsymbol{\theta})$$

$$T_2(\boldsymbol{\theta}) = \frac{\beta}{4}\sum_{i,j}(\theta_i \mathbf{X}_{\cdot i}^{\top}\mathbf{X}_{\cdot j}\theta_j)^2$$

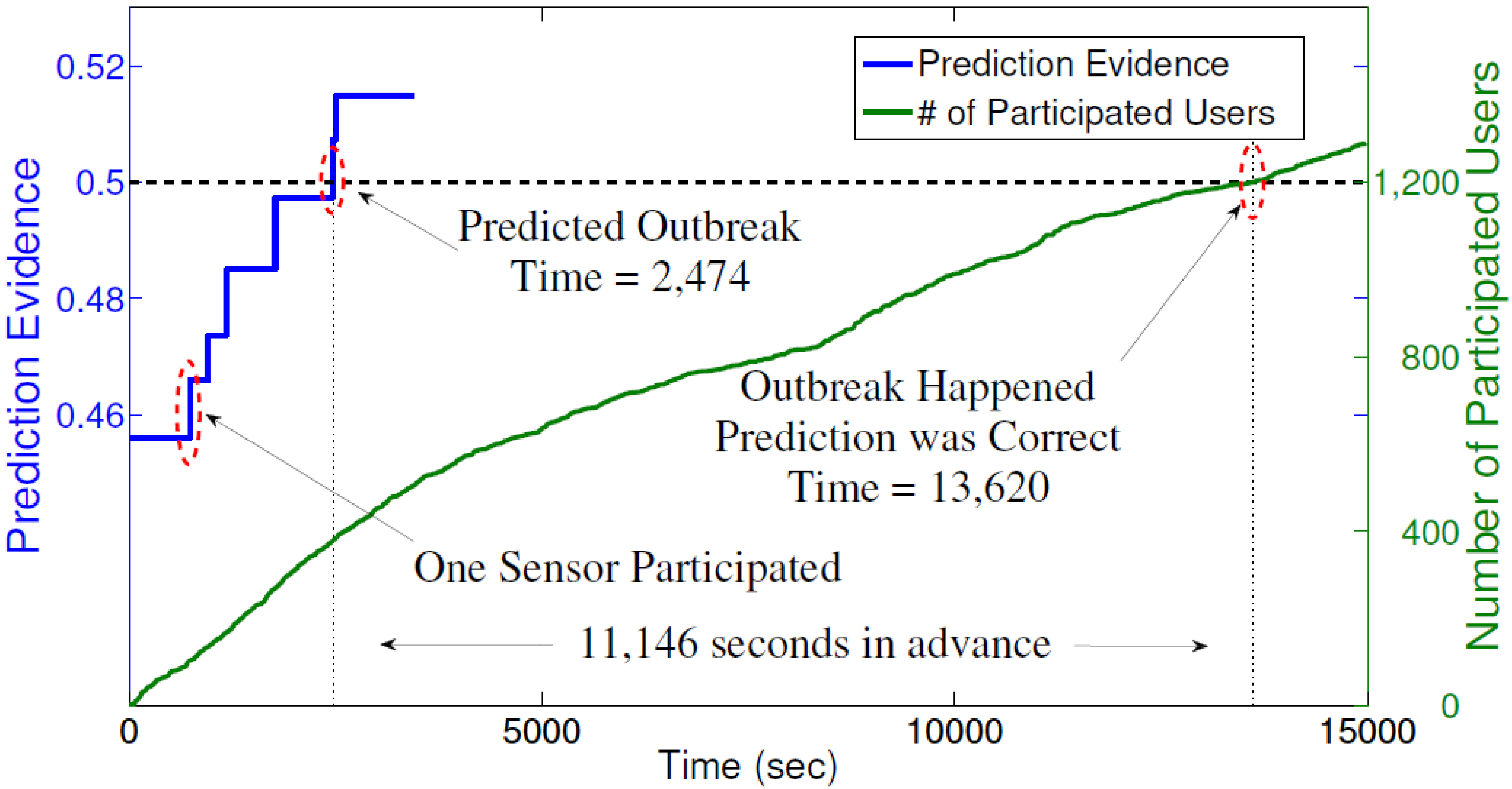
$$T_3(\boldsymbol{\theta}) = \gamma\|\boldsymbol{\theta}\|_1$$

Algorithm 1 Orthogonal Sparse LOgistic Regression (OSLOR)

Require: Tradeoff parameters $\beta > 0$, $\gamma > 0$, Radius $R > 0$, Cascade status matrix $\mathbf{X}$, Cascade outbreak indicator vector $\mathbf{y}$, Step size $c > 0$

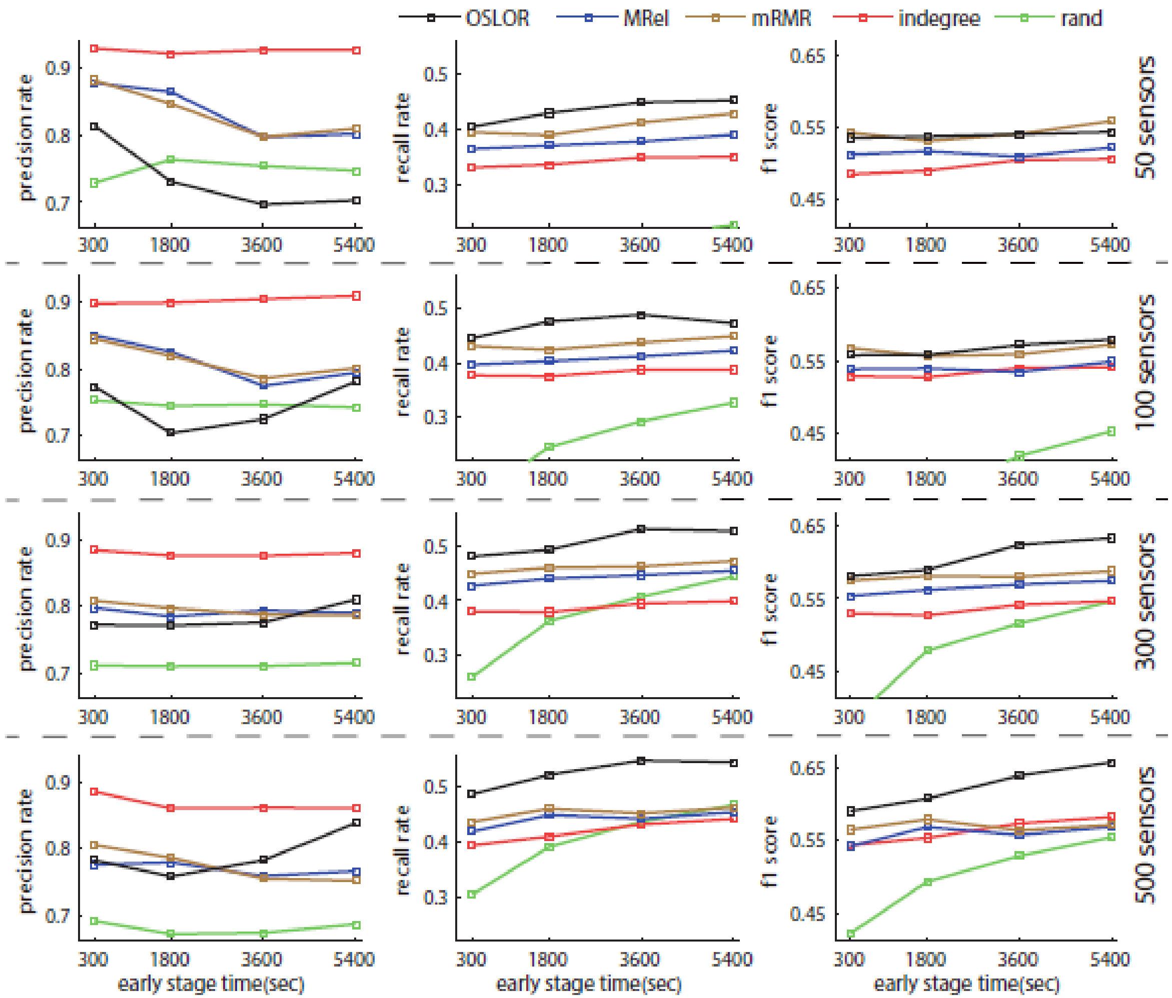
Calculate the inner product matrix $\mathbf{X}^{\top} \cdot \mathbf{X}$
Initialize the coefficient $\boldsymbol{\theta}^0 \leftarrow \mathbf{0}$
Calculate the current value of object function using Eq. (5) $F^0 \leftarrow F(\boldsymbol{\theta}^0)$
Initialize the iteration variable $k \leftarrow 0$
repeat
    Calculate gradient $\nabla g(\boldsymbol{\theta}^k)$ using Eq. (9) and Eq. (10)
    Update $\boldsymbol{\theta}^{k+1}$ using Eq. (17)
    Update the value of object function $F^{k+1} = F(\boldsymbol{\theta}^{k+1})$
    if $F^k \le F^{k+1}$ **then**
        $R \leftarrow R \cdot c$, continue;
    else
        $k \leftarrow k + 1$
    end if
until converged
Output: The final coefficient $\boldsymbol{\theta}^k$

A Showcase



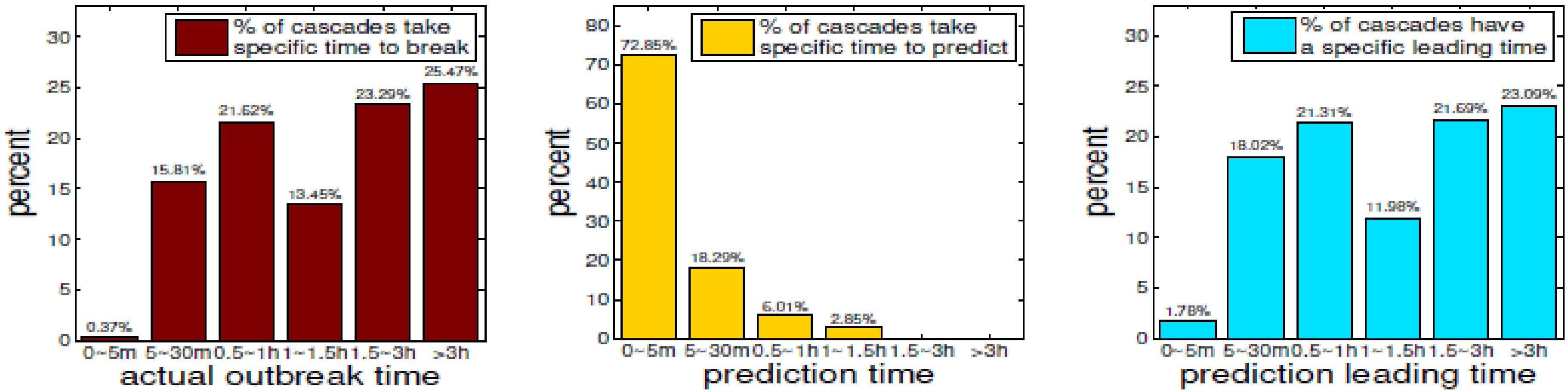
Prediction Evidence
# of Participated Users
Predicted Outbreak
Time = 2,474
Outbreak Happened
Prediction was Correct
Time = 13,620
One Sensor Participated
11,146 seconds in advance
Prediction Evidence
0.52
0.5
0.48
0.46
Number of Participated Users
1,200
800
400
0
Time (sec)
0
5000
10000
15000

Prediction Performance



- Our approach performs best
- Data driven approaches outperforms topology-based approaches
- Big nodes' participation will cause outbreaks in most cases
- Only a part of outbreaks are caused by big nodes

Prediction Leading Time



(a) (b) (c)

We only need **5 mins** **to predict the information outbreaks!**

Peng Cui, Shifei Jin, Linyun Yu, Fei Wang, Shiqiang Yang, Cascading Outbreak Prediction in Networks: A Data-Driven Approach, ***ACM SIGKDD 2013***. (Full Paper)

Discussions

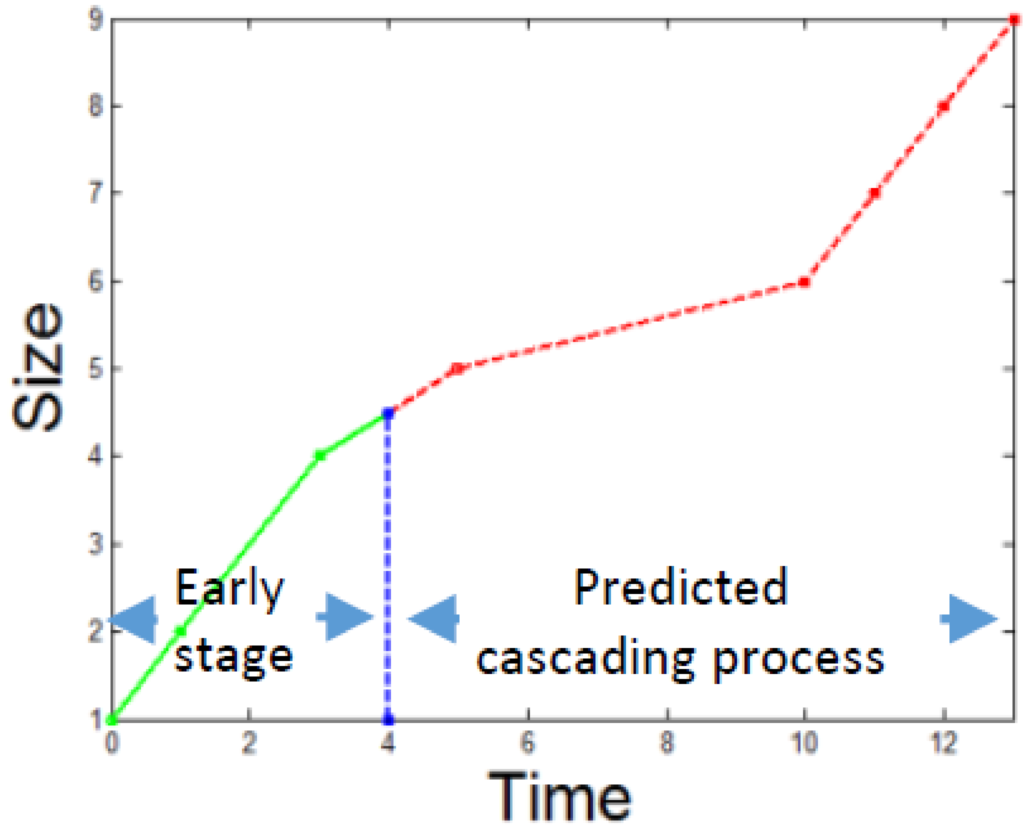
- Studying information spreading from user behavior angle is effective and promising.
- Many traditional hypothesis on the node importance and diffusion mechanism are not consistent with the real data.
- This is a one-shot study. Can we make continuous prediction on the information spreading?

Predictive Modeling on Information Spreading

Ultimate Goal: Bridge the gap between macro phenomena of information spreading and micro behavioral mechanism.

One-Hop Cascade Prediction	Cascading Outbreak Prediction	Dynamic Process Prediction
Predict the collective response of a user's followers	Predict whether the information will break out in future	Predict the dynamic cascading process of a piece of information
SIGIR'11, AAAI'11	KDD'13	ICDM'15

Beyond Cascade Size…



Time:
When will a cascade break out?

Size-Time:
How about the momentum of a cascade?

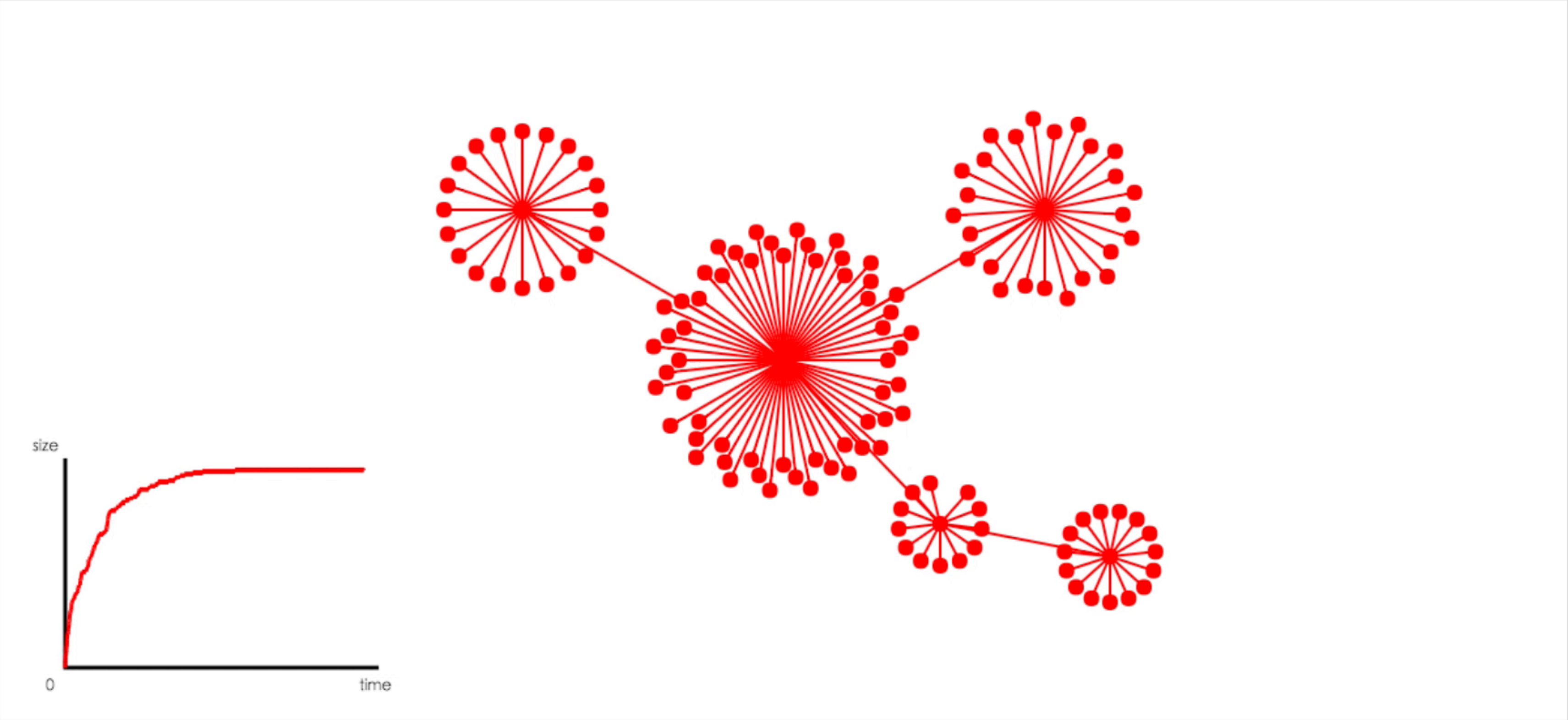
Cascading Process Prediction

Challenge: Cascade-level macro features do not work.
Content feature and structure feature
are not distinctive and predictive enough.

From Micro to Macro: Subcascades

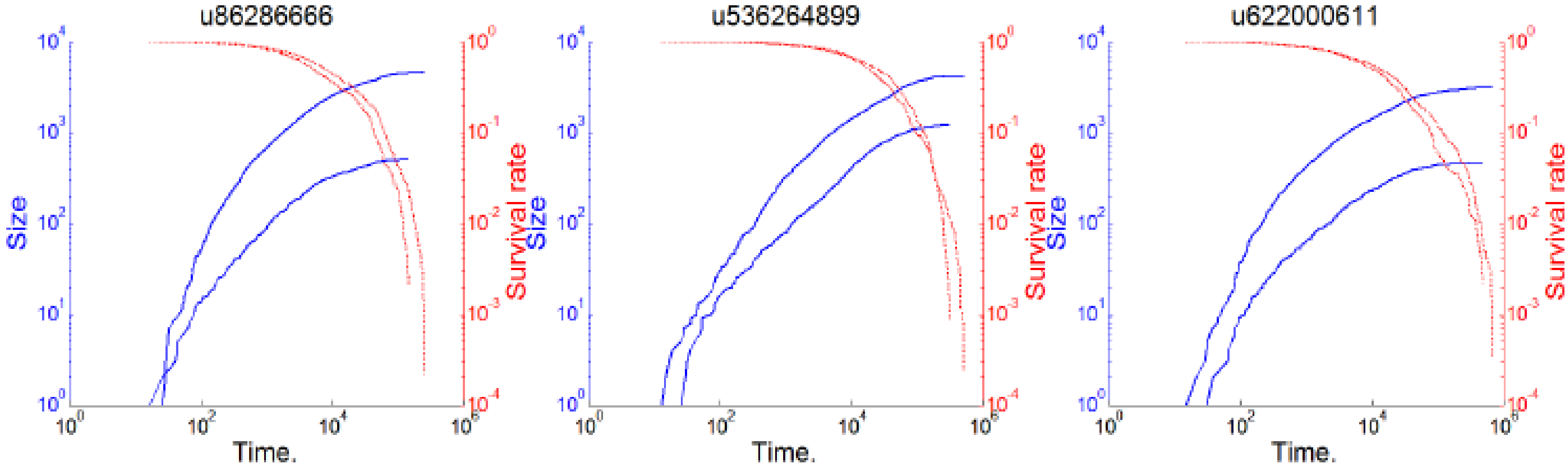
How to model subcascades?

How to connect subcascades and cascade?



Behavioral Dynamics

Behavioral Dynamics capture the changing process of the cumulative number of a user's followers retweeting a post after the user retweet the post.

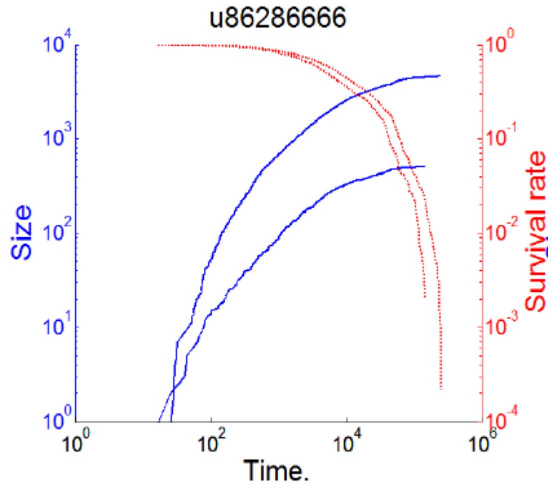


Survival Rate represent the percentage of nodes that has not been but will be infected.

Behavioral dynamics can be well represented by survival function.

Parameterize Behavioral Dynamics

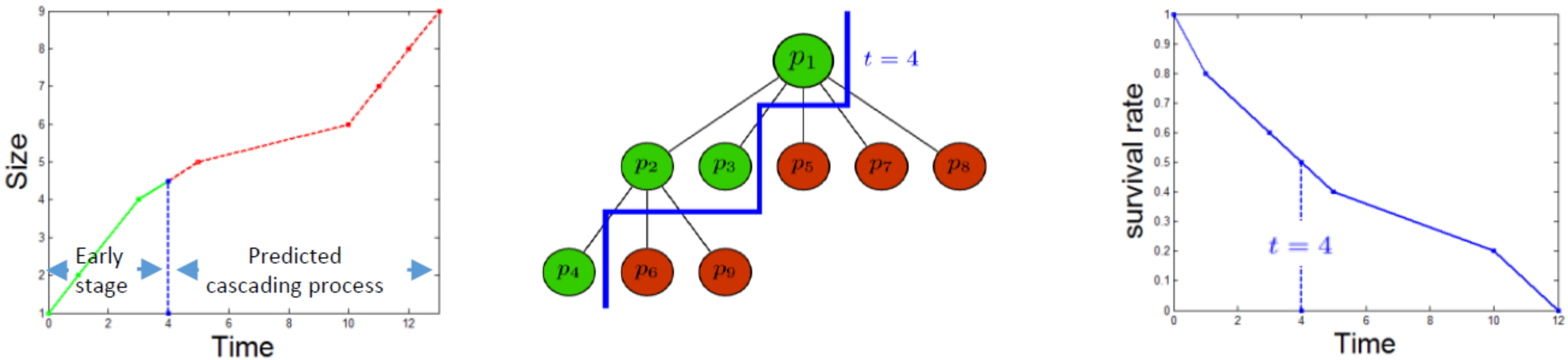
model	ks-statistic in Weibo
Exponential	0.2741
Rayleigh	0.7842
Weibull	**0.0738**



model	density function	survival function	hazard function
Exponential	$\lambda_i e^{-\lambda_i t}$	$e^{-\lambda_i t}$	λ_i
Rayleigh	$\alpha_i t e^{-\alpha_i \frac{t^2}{2}}$	$e^{-\alpha_i \frac{t^2}{2}}$	$\alpha_i t$
Weibull	$\frac{k_i}{\lambda_i}\left(\frac{t}{\lambda_i}\right)^{k_i-1} e^{-\left(\frac{t}{\lambda_i}\right)^{k_i}}$	$e^{-\left(\frac{t}{\lambda_i}\right)^{k_i}}$	$\frac{k_i}{\lambda_i}\left(\frac{t}{\lambda_i}\right)^{k_i-1}$

Characteristics of behavioral dynamics can be well captured by Weibull distribution.

From Behavioral Dynamics to Cascades



(a) Cascading Process

(b) Partially observed cascade at t

(c) Behavioral dynamics of p_1

Macro ← – – – – – – – – – – – – **Micro**

NEtworked WEibull Regression (NEWER)

$$F(\lambda, k, \beta, \gamma) = G_1(\lambda, k) + \mu G_2(\beta, \lambda) + \eta G_3(\gamma, k)$$

$$G_1(\lambda, k) = -\log L(\lambda, k)$$

$$G_2(\lambda, \beta) = \frac{1}{2N} \left\| \log \lambda - \log X \cdot \beta \right\|^2 + \alpha_\beta \left\| \beta \right\|_1$$

$$G_3(k, \gamma) = \frac{1}{2N} \left\| \log k - \log X \cdot \gamma \right\|^2 + \alpha_\gamma \left\| \gamma \right\|_1$$

- Theoretically proved to be lower-bounded.
- Coordinate Descent strategy is exploited with guaranteed convergence.

Algorithm 1 Basic Model

Input:
Set of users U involved in the cascade C before time t_{limit}, survival functions of users $S_{u_j}(t)$, predicting time t_e;

Output:
Size of cascade $size(C_{t_e})$;

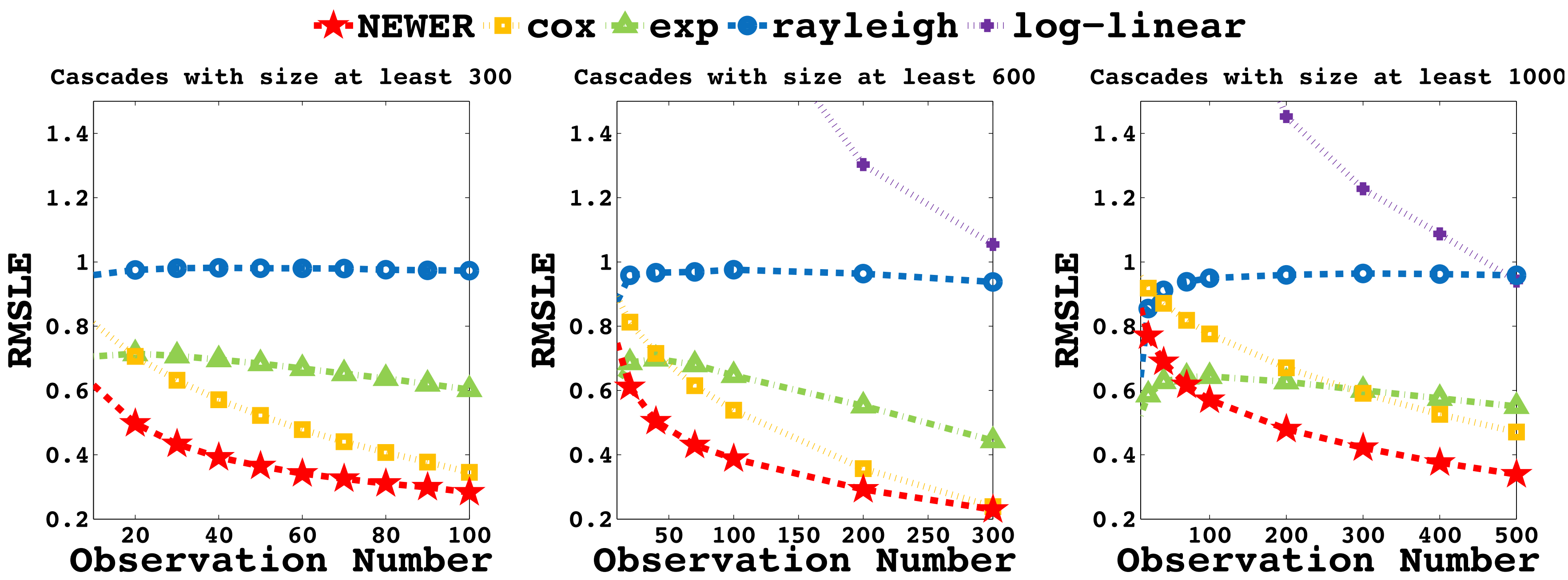
for all user $u_i \in U$ **do**
  creates a subcascade process with $replynum(u_i) = 0$
  if u_i is not root node **then**
    $replynum(rp(u_i)) = replynum(rp(u_i)) + 1$
  end if
end for
$sum = 1$
for all user $u_i \in U$ **do**
  $deathrate(u_i) = \max\left(1 - S_{u_i}(t_{limit} - t(u_i)), \frac{1}{|V|}\right)$
  $fdrate(u_i) = \max\left(1 - S_{u_i}(t_e - t(u_i)), \frac{1}{|V|}\right)$
  $sum = sum + \frac{replynum(u_i) \cdot fdrate(u_i)}{deathrate(u_i)}$
end for
return $size(C_{t_e}) = sum$

Experiments

- Datasets: Tencent Weibo
  - All cascades generated between Nov 15th and Nov 25th in 2011.
  - retain all 0.59 million cascades that the cascades size are at least 5.
- Baseline:
  - Cox Proportional Hazard Regression Model (Cox)
  - Exponential/Rayleigh Proportional Hazard Regression Model (Exponential/Rayleigh)
  - log-Linear regression(Log-linear)
- Evaluation metric:
  - RMSLE: Root Mean Square Log Error
  - $\Delta\sigma$-Precision: Precision value that the predicted value within $(1+\sigma)\pm1$groundtruth

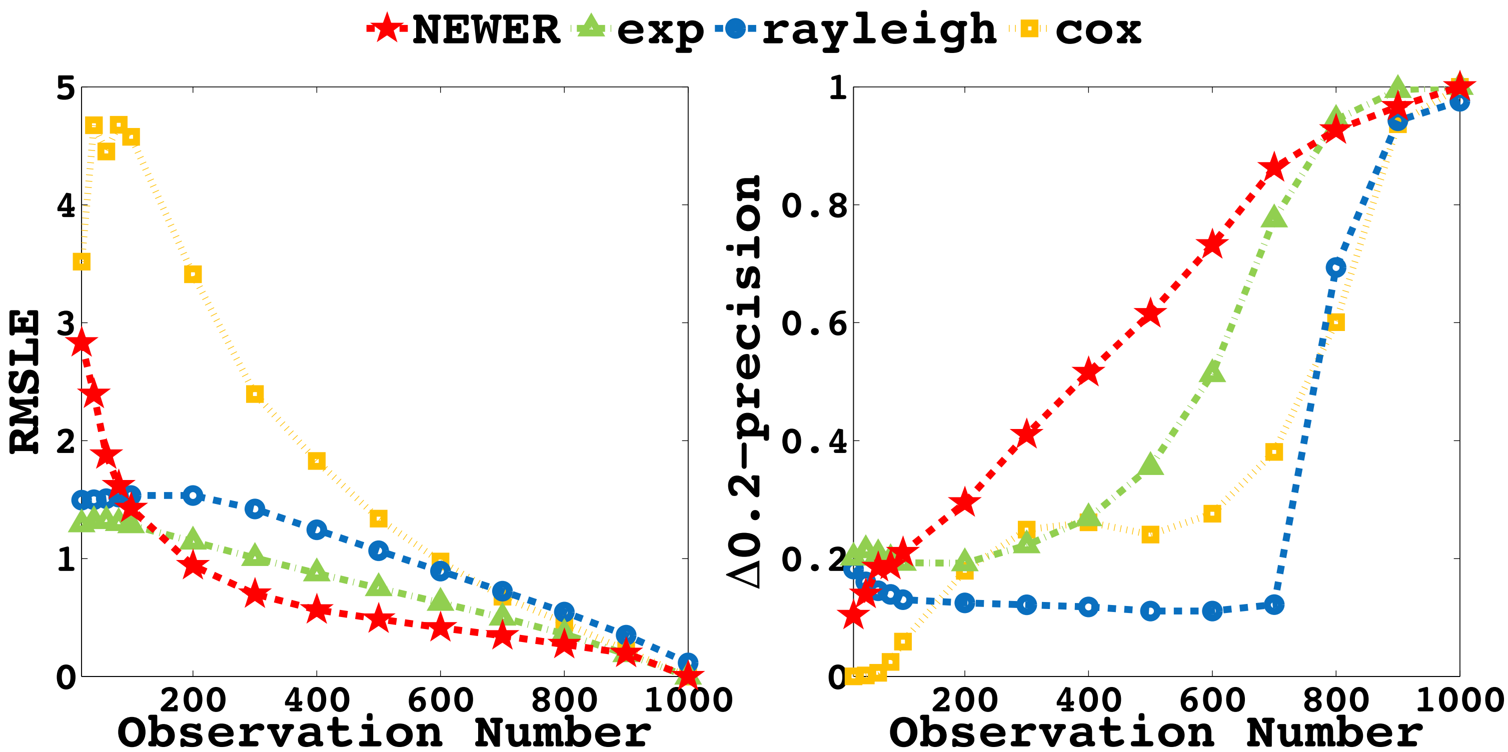
Cascade Size Prediction

- What is the final size of the cascade?



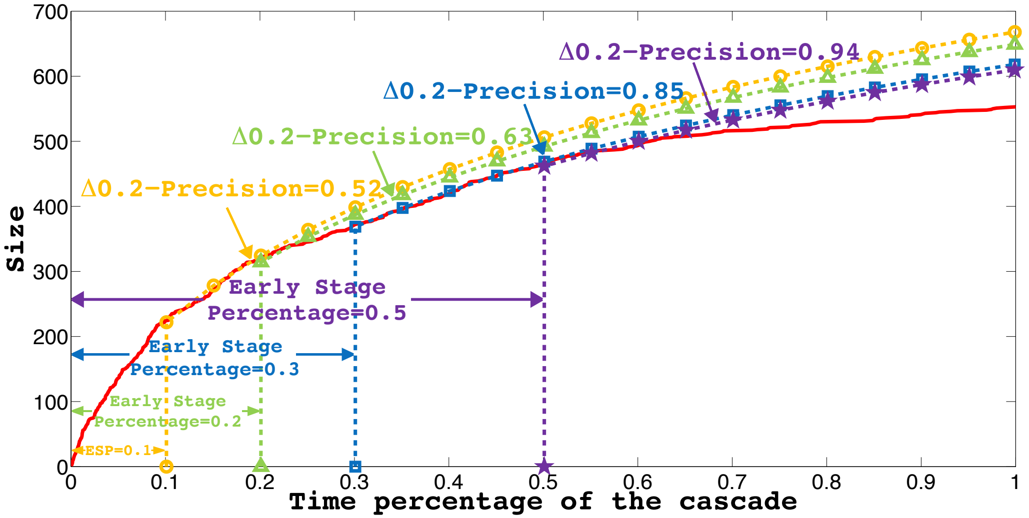
Outbreak Time Prediction

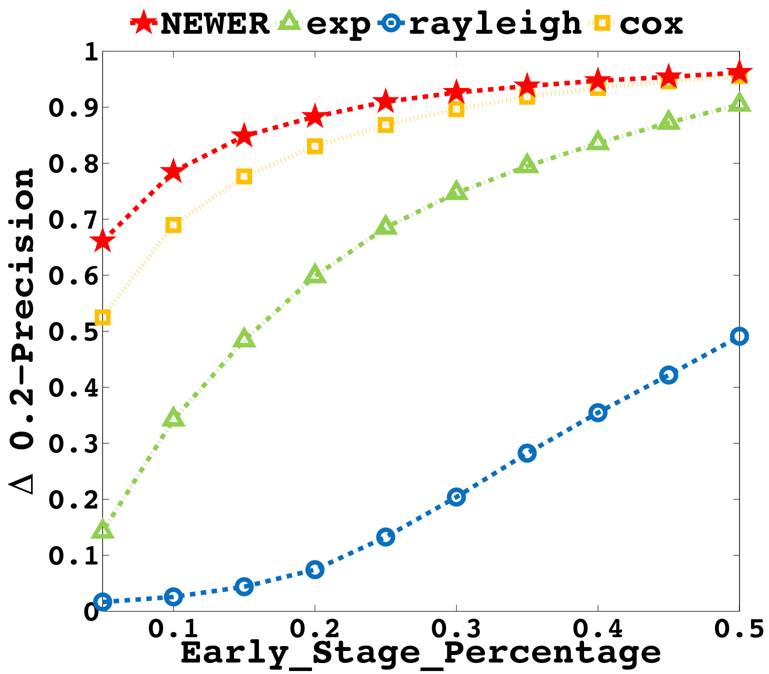
- When will the cascade break out?



Cascading Process Prediction

- What is the size of the cascade at any later point?





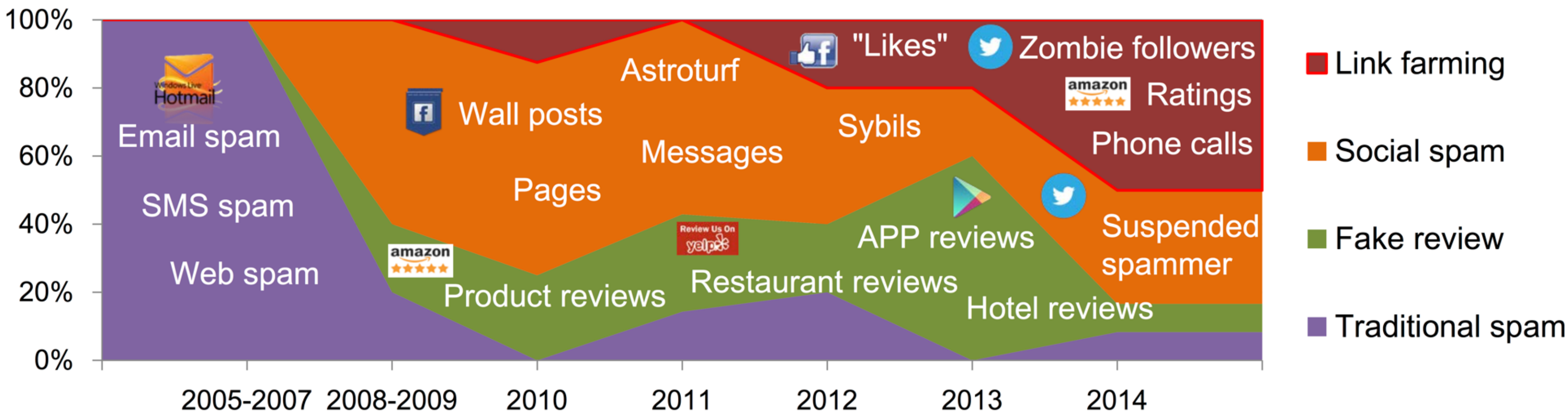
Conclusions

- Before predicting information spreading, understanding the ***behavioral mechanism*** is critical and fundamental.
- Behaviors can be modeled in different ***granularities***, which depends on the target problem.
- Modeling information spreading with ***continuous-time model*** is promising and demonstrated to be effective in our research.

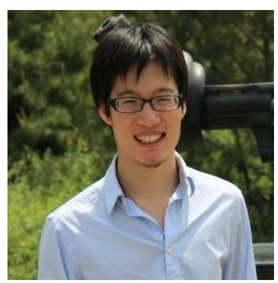
Outline

- Prediction for natural behavior
  - Modeling individual behavior (MICRO)
  - Modeling information cascade (MACRO)
- **Detection for unnatural behavior**
  - **Suspicious behavior detection**

Suspicious Behavior Detection



JIANG, Meng (UIUC)



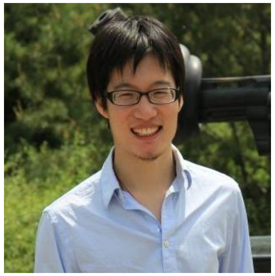
CUI, Peng (Tsinghua)

FALOUTSOS, Christos (CMU)

Suspicious Behavior Detection: Current Trends and Future Directions. *Special Issue on Online Behavioral Analysis and Modeling, IEEE Intelligent Systems Magazine (ISSI)*, 2016. (to appear)

Suspicious Behavior Detection

JIANG, Meng (UIUC)



CUI, Peng (Tsinghua)

FALOUTSOS, Christos (CMU)

Suspicious Behavior Detection: Current Trends and Future Directions. *Special Issue on Online Behavioral Analysis and Modeling, IEEE Intelligent Systems Magazine (ISSI)*, 2016. (to appear)

Social Link Farming

- Selling Twitter followers

Social Link Farming

- Selling Facebook Likes

25,000 Facebook Likes	50,000 Facebook Likes	100,000 Facebook Likes	200,000 Facebook Likes
$265	$525	$1,000	$1,750
Lifetime Replacement Warranty	Lifetime Replacement Warranty	Lifetime Replacement Warranty	Lifetime Replacement Warranty
Dedicated 24/7 Customer Service	Dedicated 24/7 Customer Service	Dedicated 24/7 Customer Service	Dedicated 24/7 Customer Service
100% Risk Free, Try Us Today	100% Risk Free, Try Us Today	100% Risk Free, Try Us Today	100% Risk Free, Try Us Today
Order starts within 24 - 48 hours	Order starts within 24 - 48 hours	Order starts within 24 -48 hours	Order starts within 24 -48 hours
Order completed within 22 days	Order completed within 35 days	Order completed within 35 days	Order completed within 35 days

Suspicious Behavior Detection

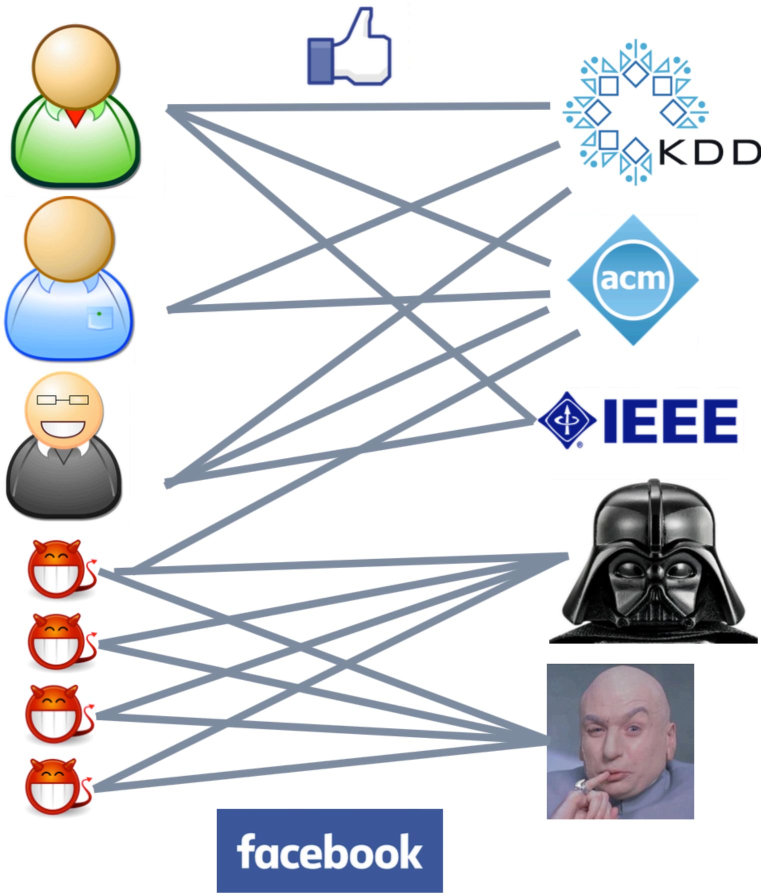
- Detecting suspicious behavioral patterns
  - **Spotting:** Lockstep patterns
    - Methods that can spot strange behaviors
  - **Catching:** Synchronized patterns
    - Scalable algorithms with theoretical guarantee
  - **Solving:** Suspiciousness in multiple dimensions
    - A principled metric for suspiciousness

Lockstep Behavior: Facebook Likes

25,000 Facebook Likes	50,000 Facebook Likes	100,000 Facebook Likes	200,000 Facebook Likes
$265	$525	$1,000	$1,750
Lifetime Replacement Warranty	Lifetime Replacement Warranty	Lifetime Replacement Warranty	Lifetime Replacement Warranty
Dedicated 24/7 Customer Service	Dedicated 24/7 Customer Service	Dedicated 24/7 Customer Service	Dedicated 24/7 Customer Service
100% Risk Free, Try Us Today	100% Risk Free, Try Us Today	100% Risk Free, Try Us Today	100% Risk Free, Try Us Today
Order starts within 24 - 48 hours	Order starts within 24 - 48 hours	Order starts within 24 -48 hours	Order starts within 24 -48 hours
Order completed within 22 days	Order completed within 35 days	Order completed within 35 days	Order completed within 35 days

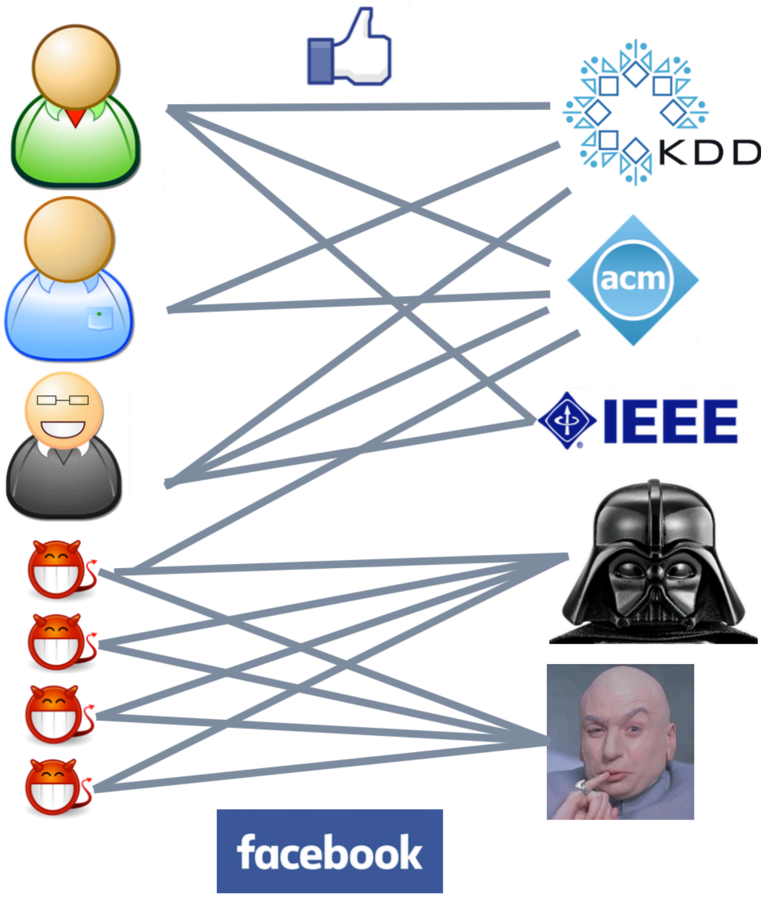
Beutel et al. **CopyCatch: Stopping Group Attacks by Spotting Lockstep Behavior in Social Neworks.** *WWW*, 2013.

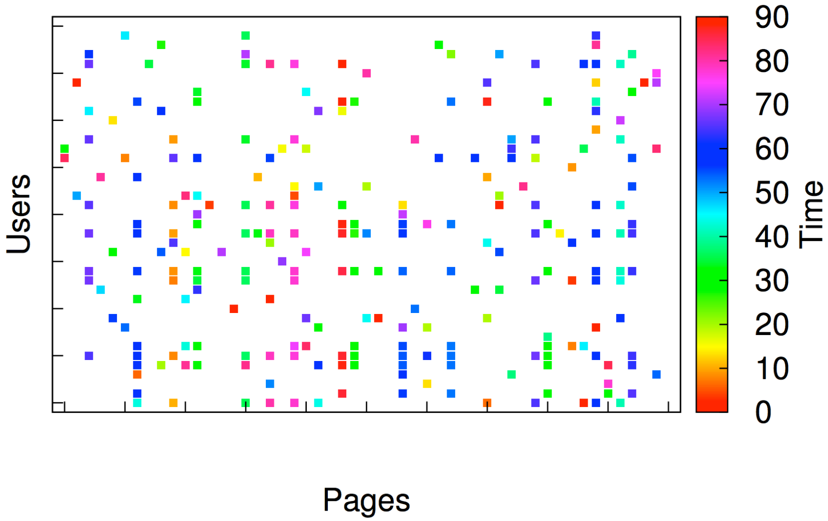
Lockstep Behavior: Facebook Likes



Beutel et al. **CopyCatch: Stopping Group Attacks by Spotting Lockstep Behavior in Social Neworks.** *WWW*, 2013.

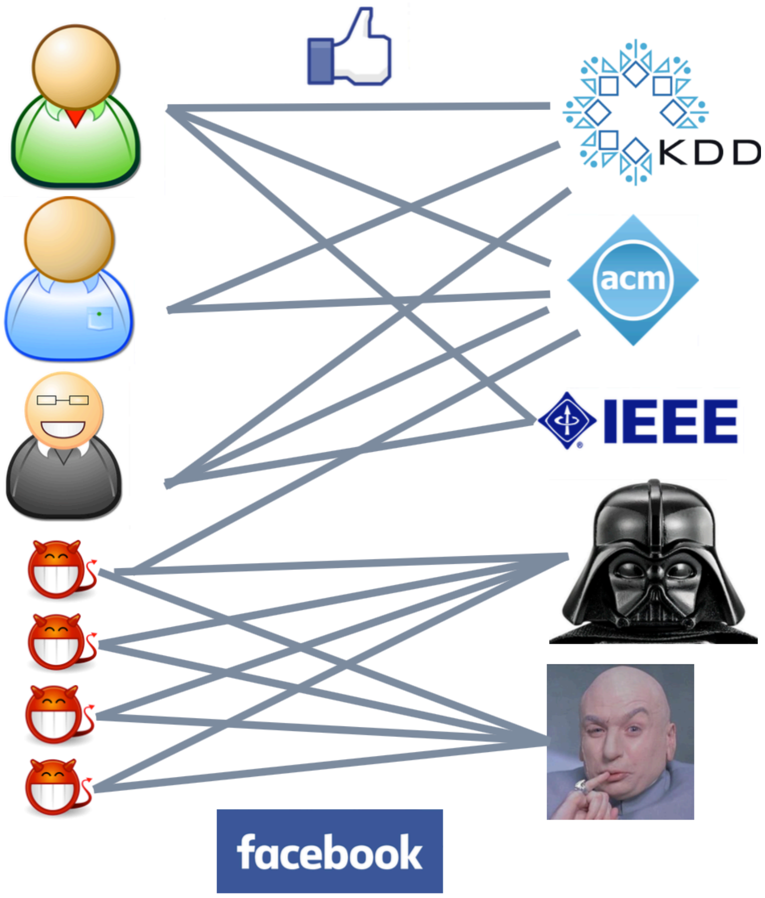
Lockstep Behavior: Graphical View

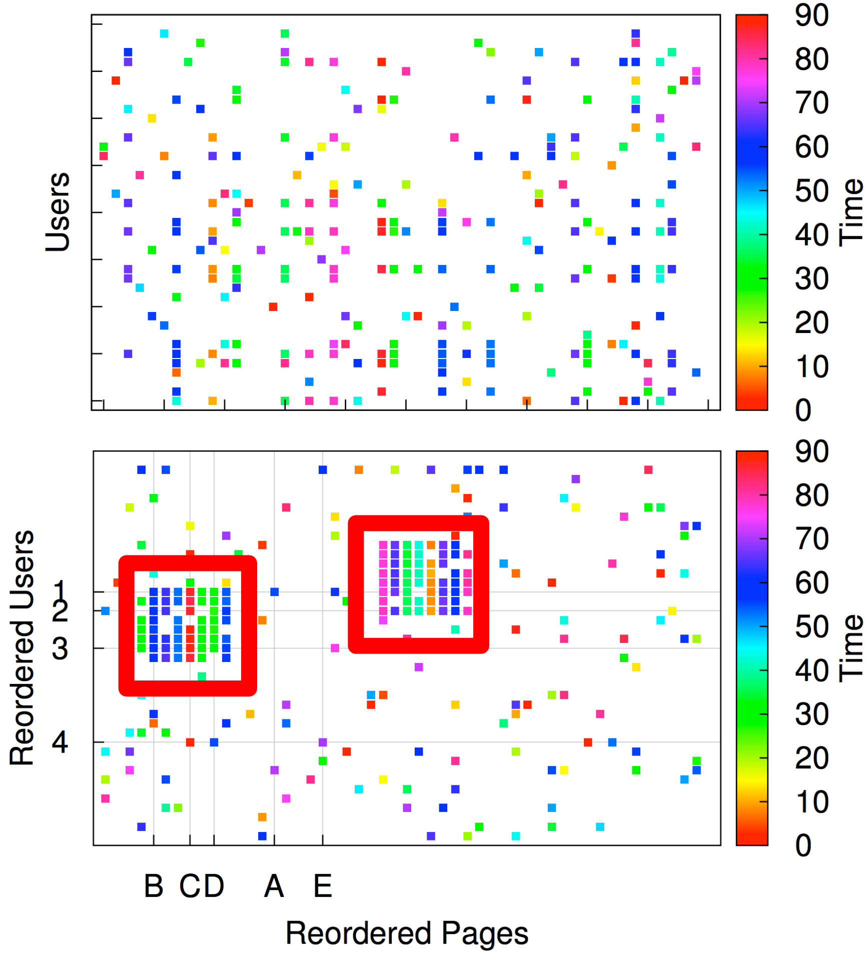




Beutel et al. **CopyCatch: Stopping Group Attacks by Spotting Lockstep Behavior in Social Neworks.** *WWW*, 2013.

Lockstep Behavior: Reorder Matrix

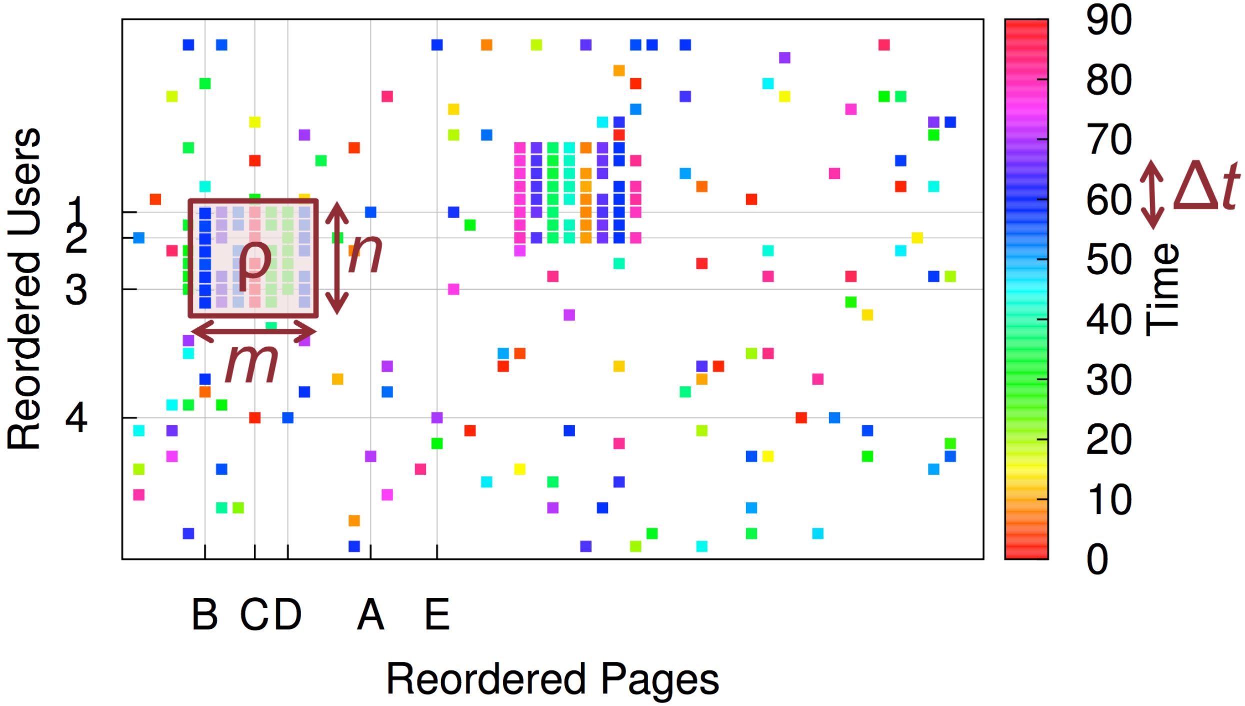




Beutel et al. **CopyCatch: Stopping Group Attacks by Spotting Lockstep Behavior in Social Neworks.** *WWW*, 2013.

Lockstep Behavior: Seed + Search

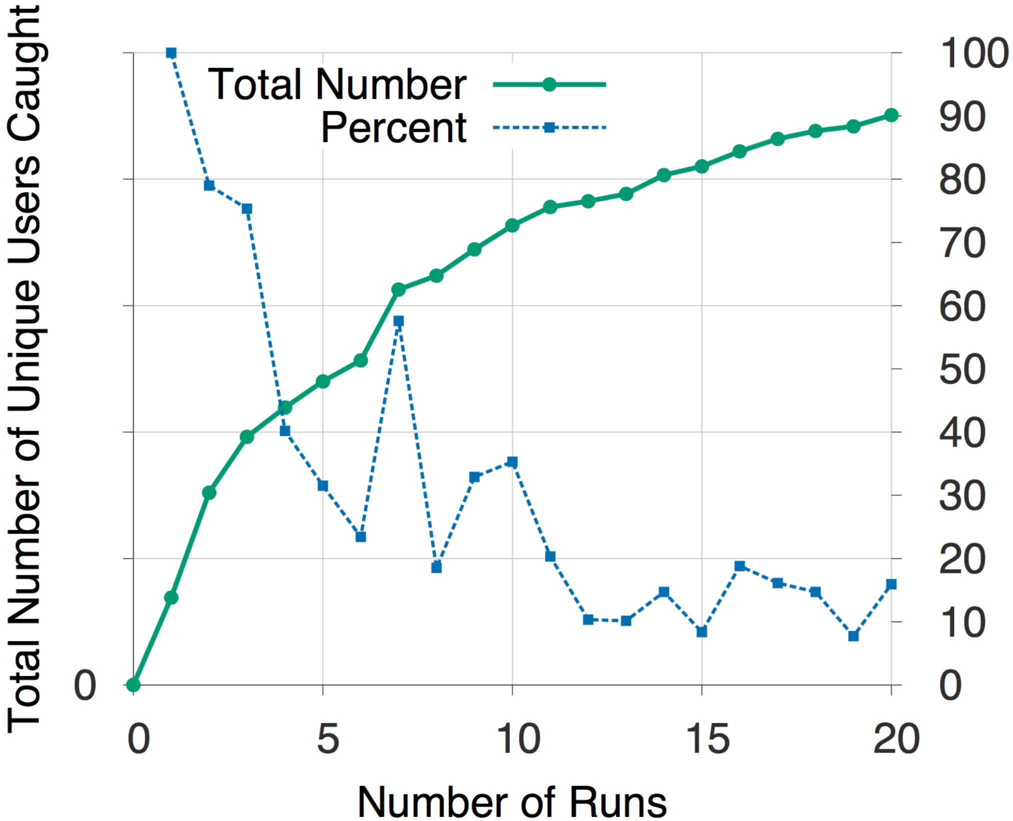
- CopyCatch
  - "Near Bipartite Core": n users, m Pages, ρ, Δt



Beutel et al. **CopyCatch: Stopping Group Attacks by Spotting Lockstep Behavior in Social Neworks.** *WWW*, 2013.

Lockstep Behavior: Performance

- CopyCatch

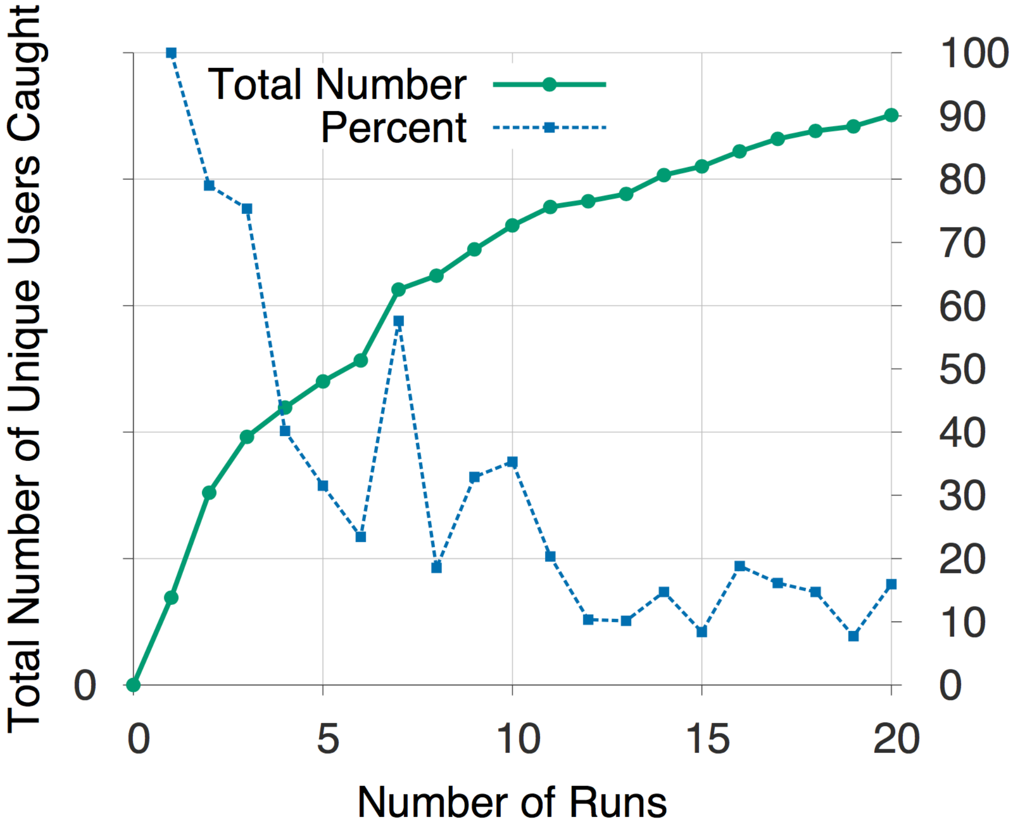


Beutel et al. **CopyCatch: Stopping Group Attacks by Spotting Lockstep Behavior in Social Neworks.** *WWW*, 2013.

Lockstep Behavior: Performance

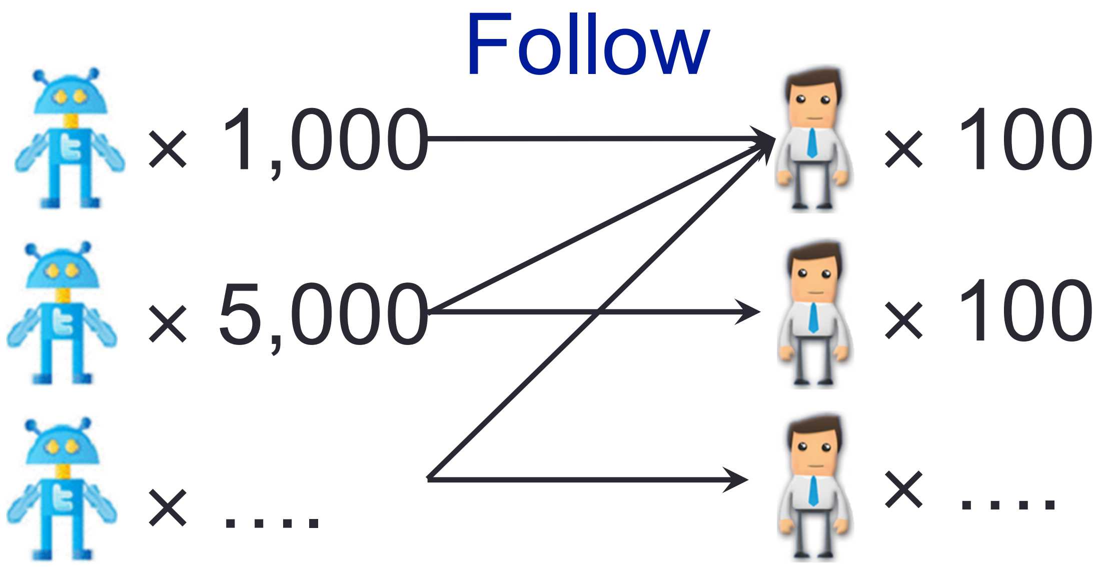
❖CopyCatch

Seed Selection!!!



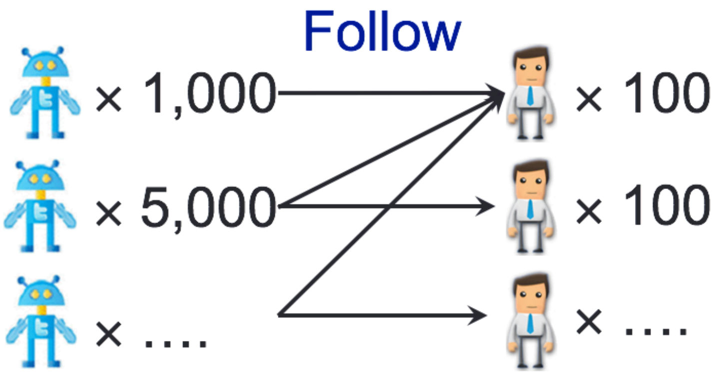
Beutel et al. **CopyCatch: Stopping Group Attacks by Spotting Lockstep Behavior in Social Neworks.** *WWW*, 2013.

Lockstep Behavior: Twitter Followers



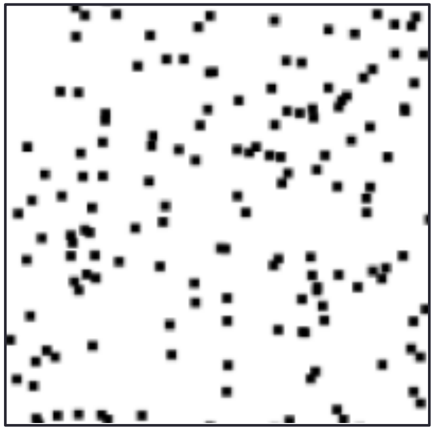
Jiang et al. **Inferring Strange Behavior from Connectivity Pattern in Social Networks.** *PAKDD*, 2014.

Lockstep Behavior: Reorder Matrix

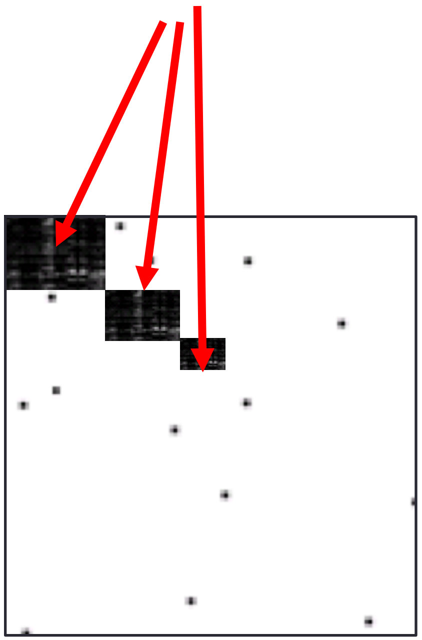


Seeds

Given

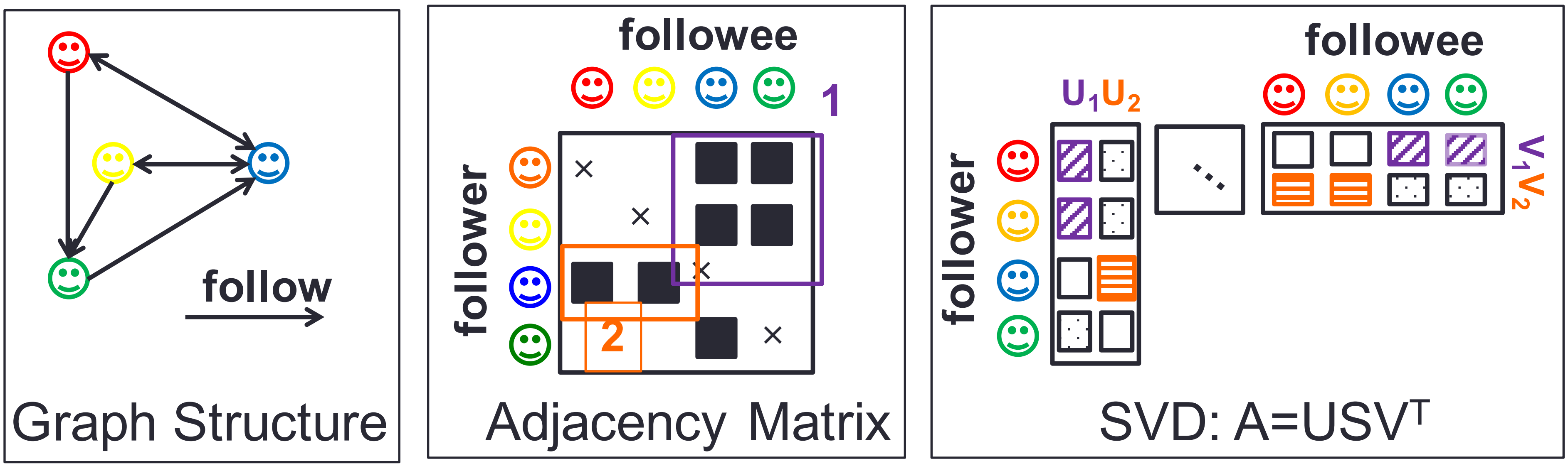


Reorder



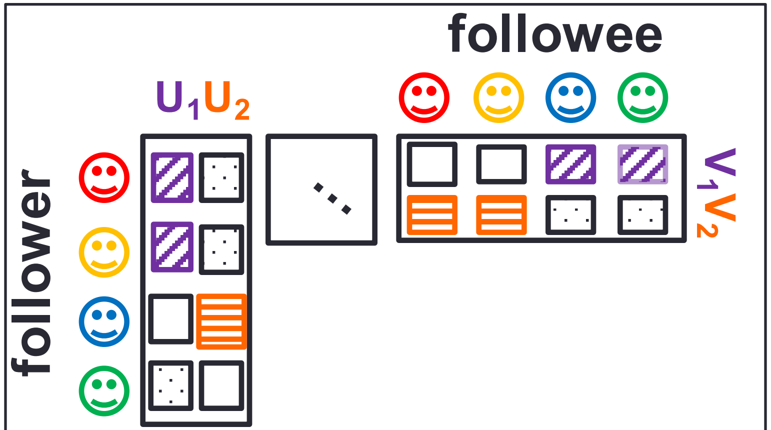
Jiang et al. **Inferring Strange Behavior from Connectivity Pattern in Social Networks.** *PAKDD*, 2014.

Lockstep Behavior: SVD Reminder



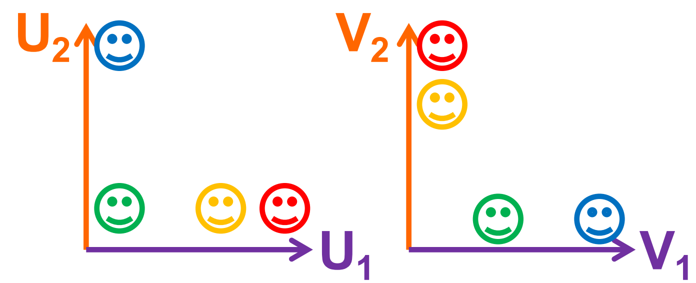
Jiang et al. **Inferring Strange Behavior from Connectivity Pattern in Social Networks.** *PAKDD*, 2014.

Lockstep Behavior: Spectral Subspace



SVD: $A=USV^T$

Pairs of singular vectors:

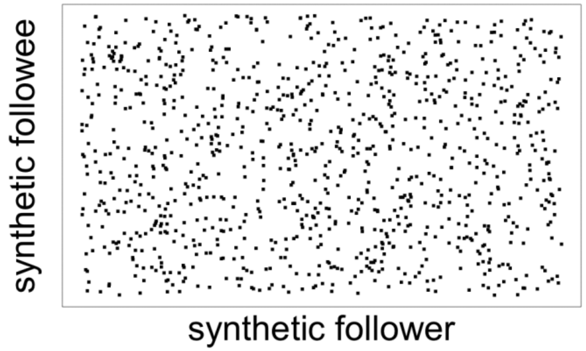


"Spectral Subspace Plot"

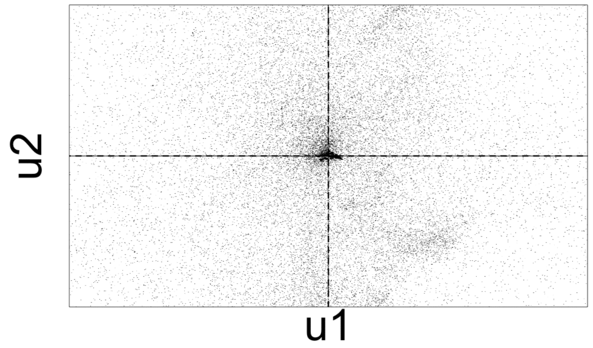
Spectral Subspace Plot: Case #0

- NO lockstep behavior: Scatter

Adjacency Matrix



Spectral Subspace Plot



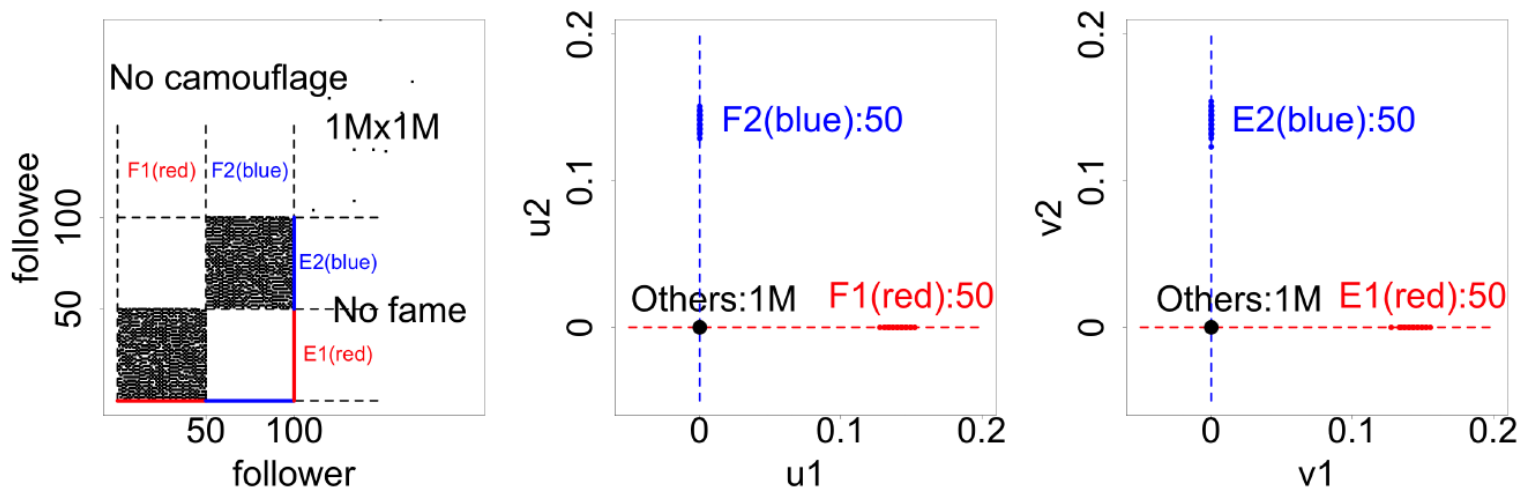
Jiang et al. **Inferring Strange Behavior from Connectivity Pattern in Social Networks.** *PAKDD*, 2014.

Spectral Subspace Plot: Case #1

❖Non-overlapping lockstep: "Rays"

Adjacency Matrix

Spectral Subspace Plot

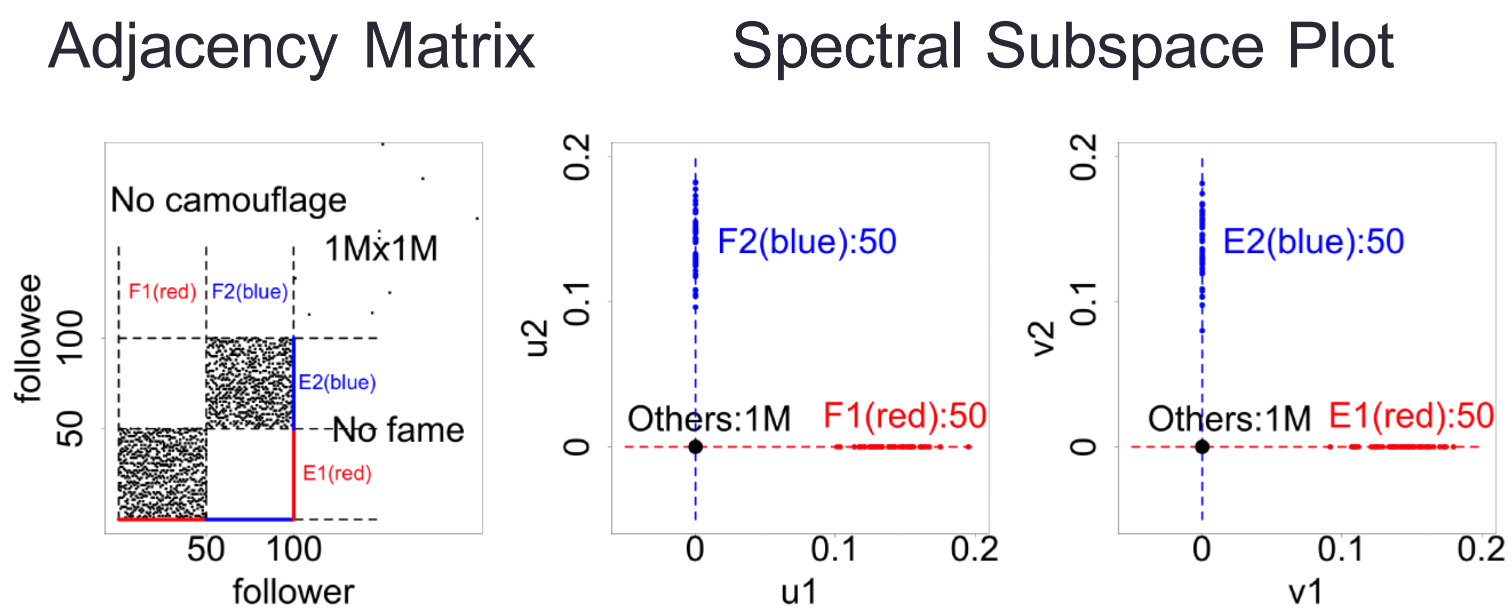


Rule 1 (short "rays"): two blocks, high density (90%), no "camouflage", no "fame"

Jiang et al. **Inferring Strange Behavior from Connectivity Pattern in Social Networks.** *PAKDD*, 2014.

Spectral Subspace Plot: Case #2

- Non-overlapping: Low density, Elongation

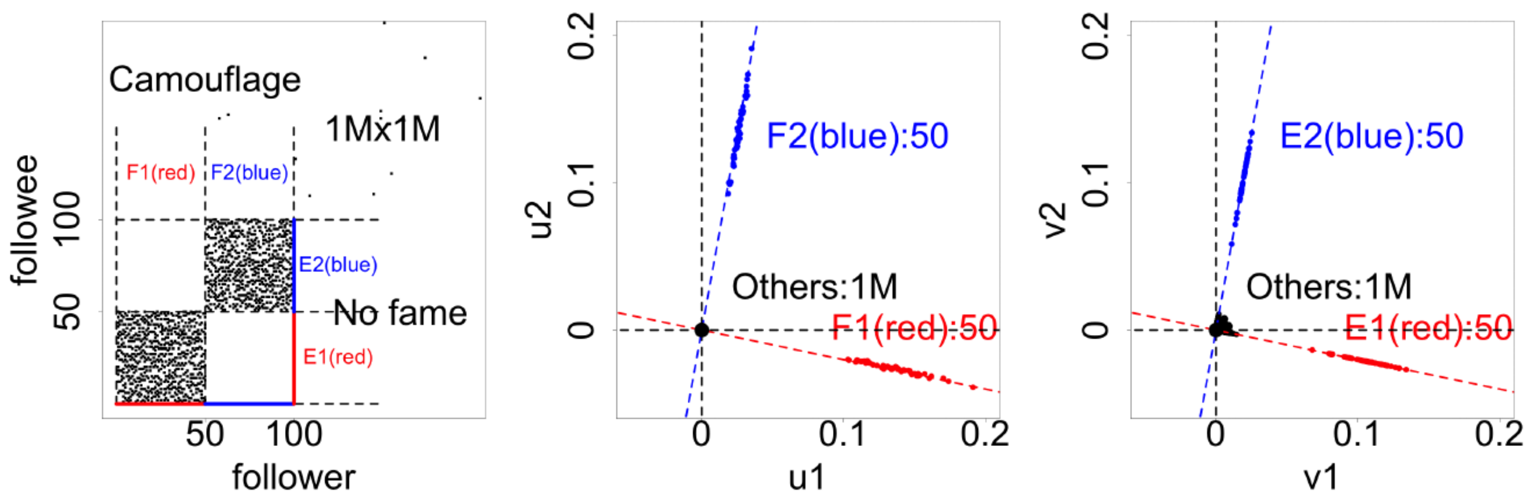


Rule 2 (long "rays"): two blocks, low density (50%), no "camouflage", no "fame"

Spectral Subspace Plot: Case #3

- Non-overlapping: Camouflage/Fame, Tilting

Adjacency Matrix   Spectral Subspace Plot

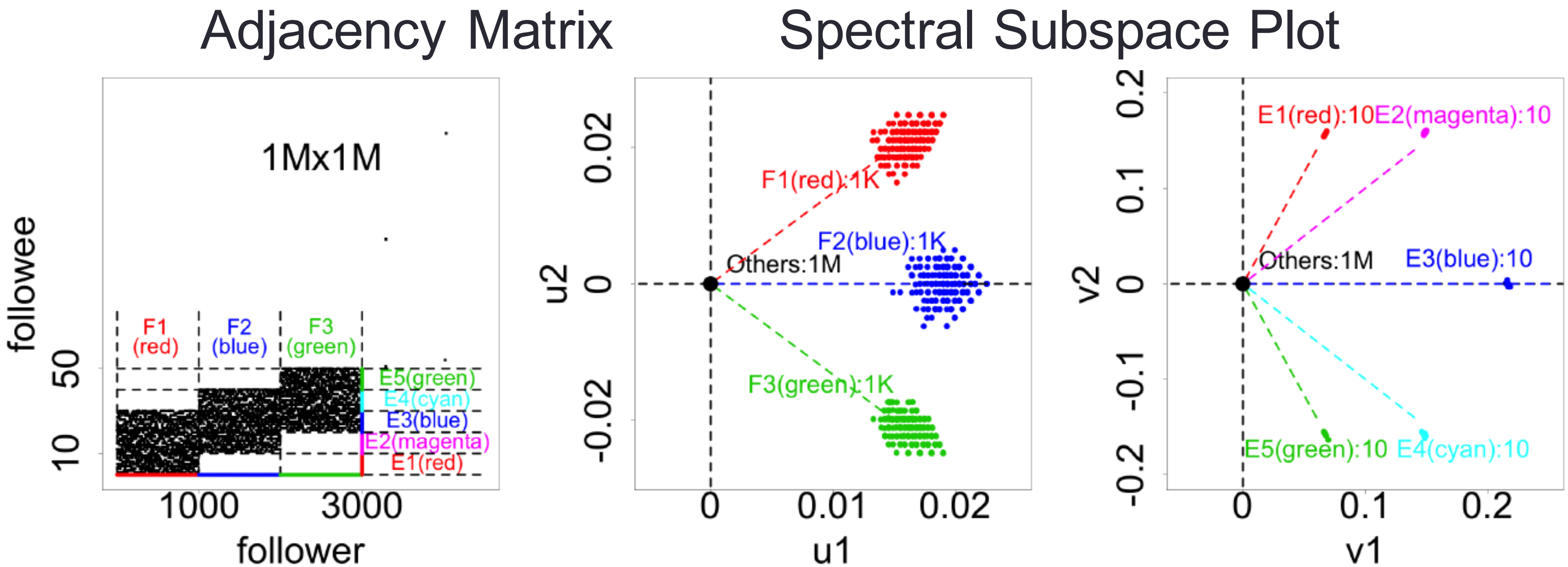


Rule 3 (tilting "rays"): two blocks, with "camouflage", no "fame"

Jiang et al. **Inferring Strange Behavior from Connectivity Pattern in Social Networks.** *PAKDD*, 2014.

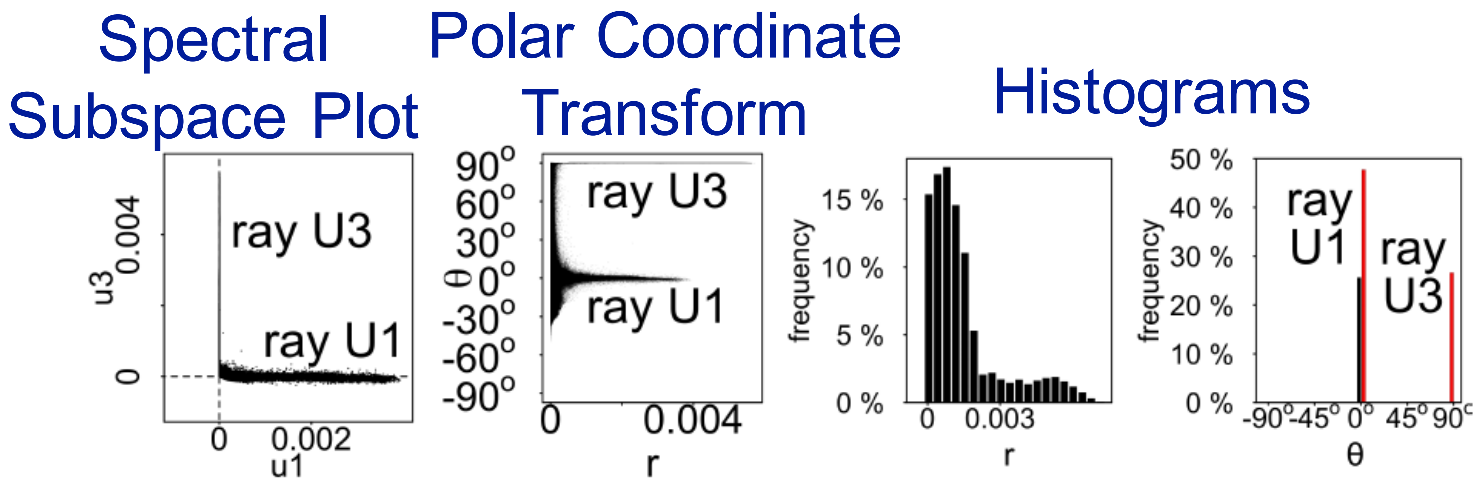
Spectral Subspace Plot: Case #4

- Overlapping: "Staircase", "Pearls"

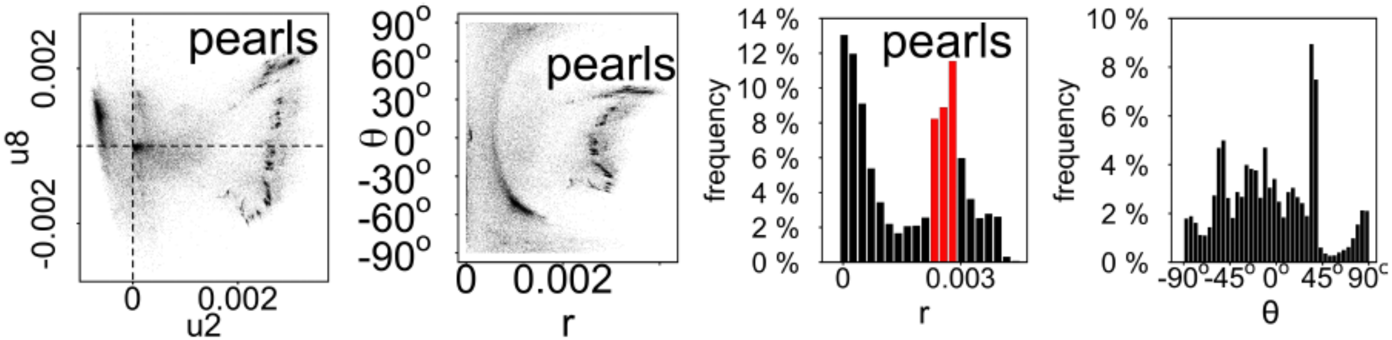


Rule 4 ("pearls"): a "staircase" of three partially overlapping blocks.

Spectral Subspace Plot: Reading & LockInfer



"rays" show two apparent spikes on θ frequency at 0^o and 90^o

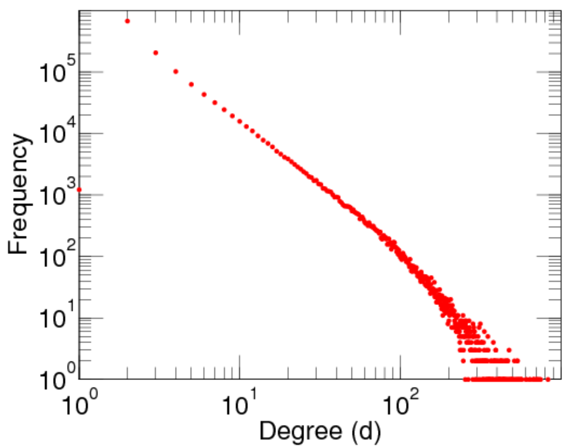
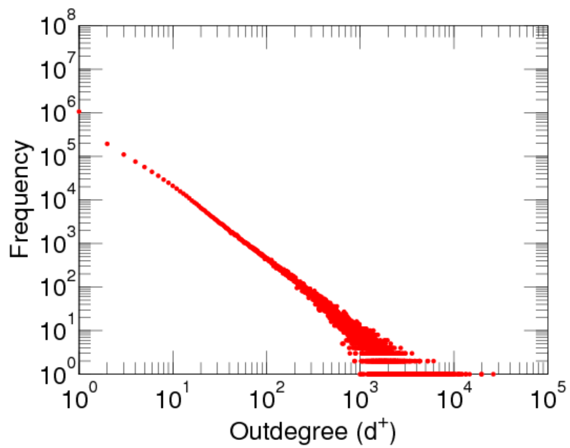
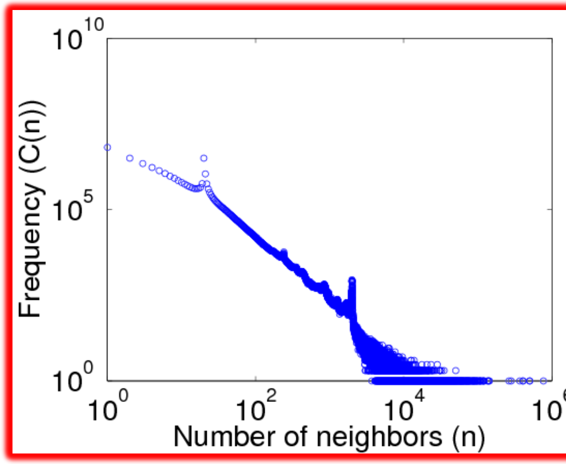


"pearls" show a spike on r frequency at a much-greater-than-zero value

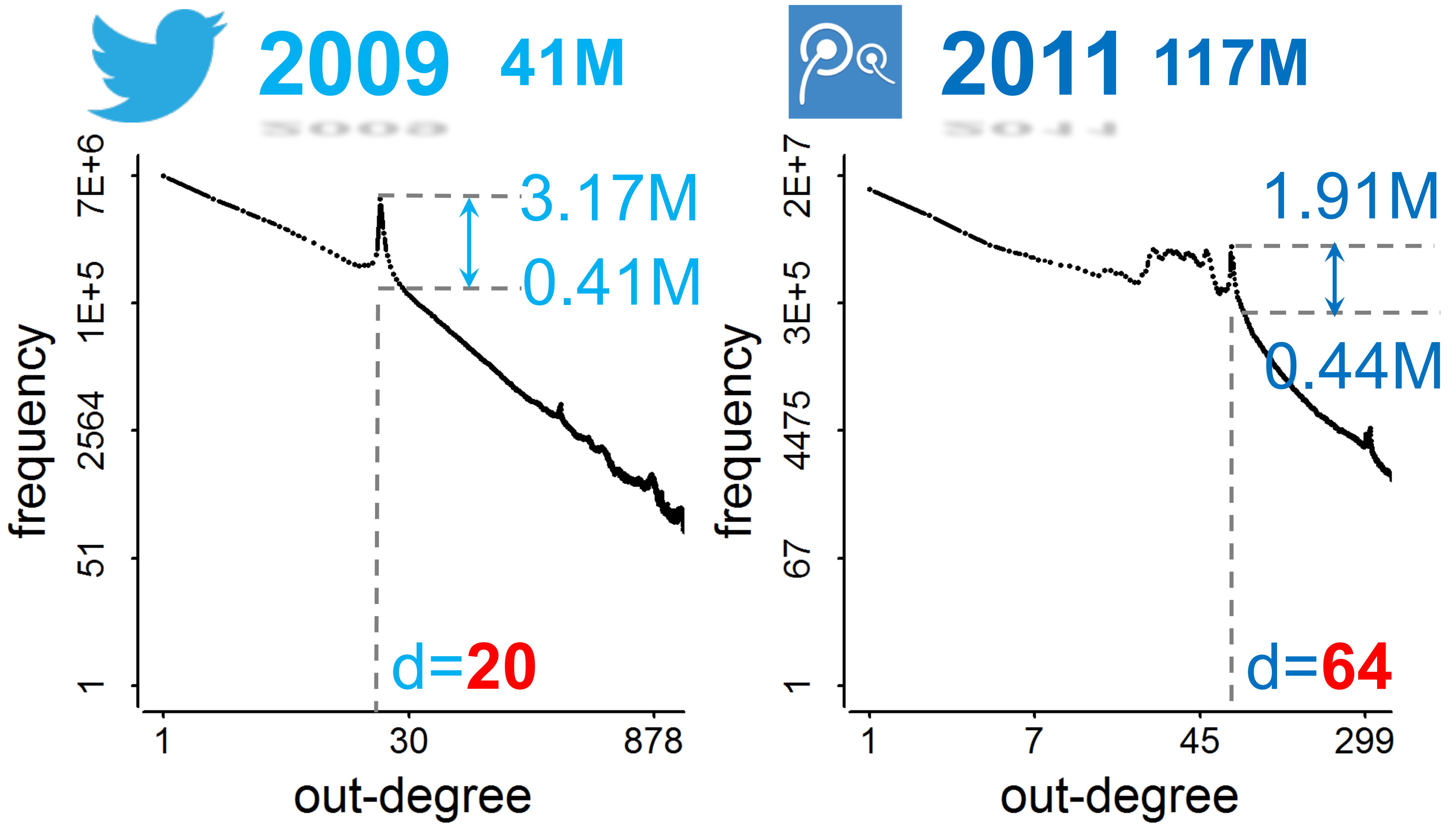
Jiang et al. **Inferring Strange Behavior from Connectivity Pattern in Social Networks.** *PAKDD*, 2014.

Graphical View: Power-Law Distribution

- Out-degree distribution

DBLP **Author**-publication	Flickr **User**-user	Twitter **Who**-follows-whom
		

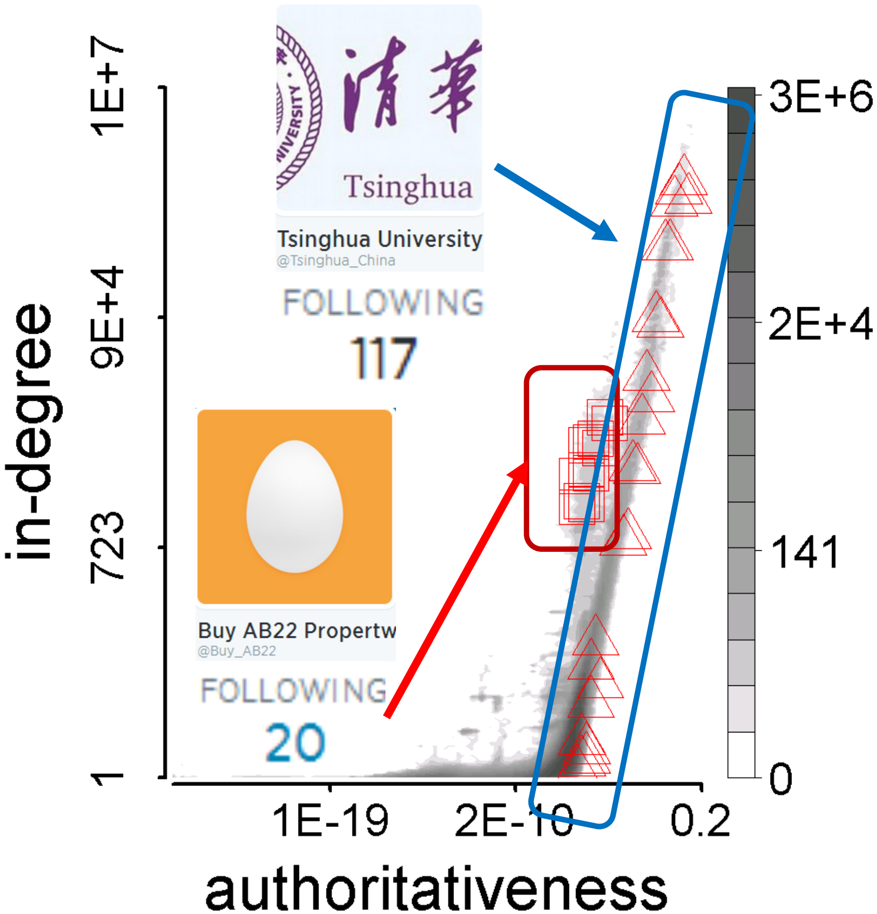
[konect.uni-koblenz.de/networks/]



Jiang et al. **CatchSync: Catching Synchronized Behavior in Large Directed Graphs.** *KDD Best Paper Finalist*, 2014.

Synchronized Behavior: Features

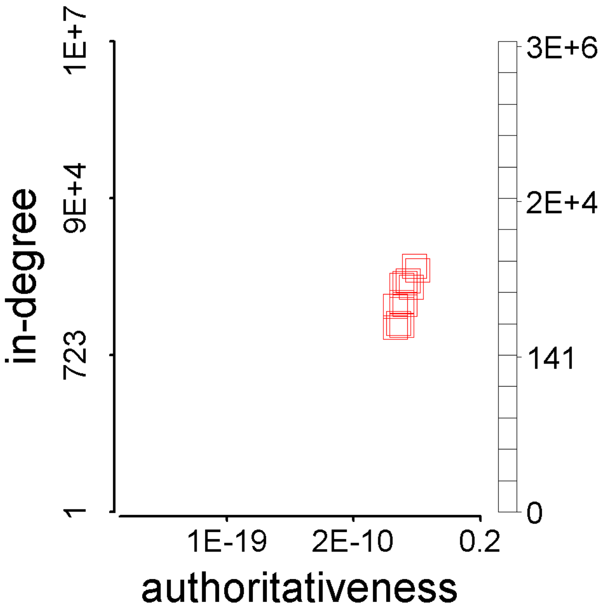
- ❖Synchronized
- ❖Abnormal

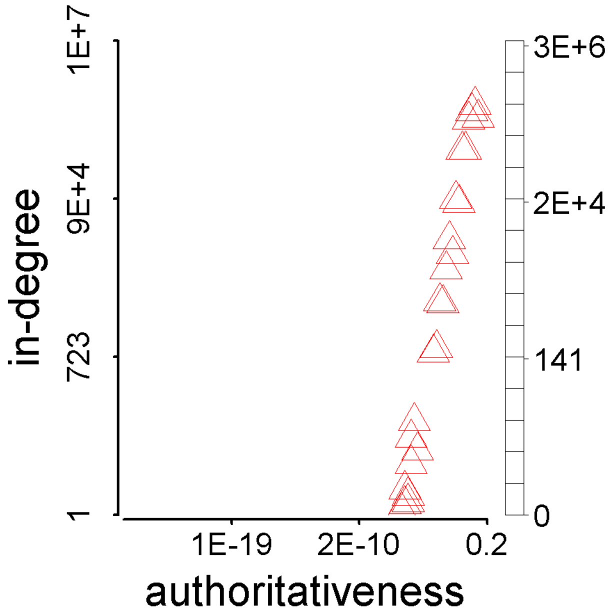


Jiang et al. **CatchSync: Catching Synchronized Behavior in Large Directed Graphs.** *KDD Best Paper Finalist*, 2014.

Synchronized Behavior: Synchronicity

$$sync(u) = \frac{\sum_{(v,v')\in\mathcal{F}(u)\times\mathcal{F}(u)} \mathbf{p}(v)\cdot\mathbf{p}(v')}{d(u)\times d(u)}$$

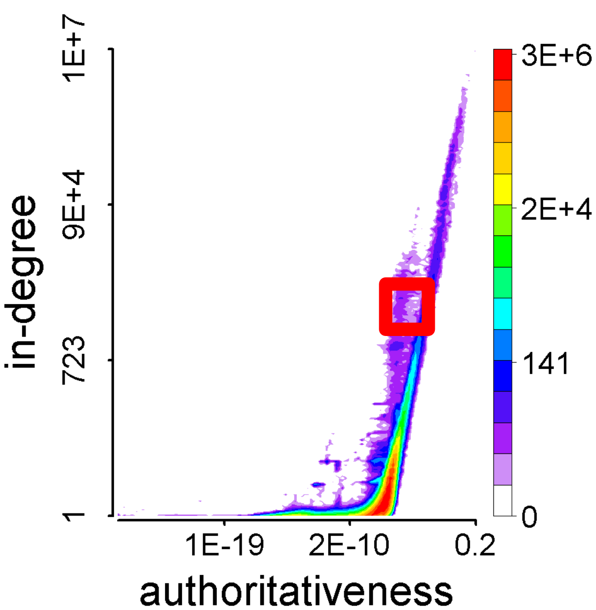


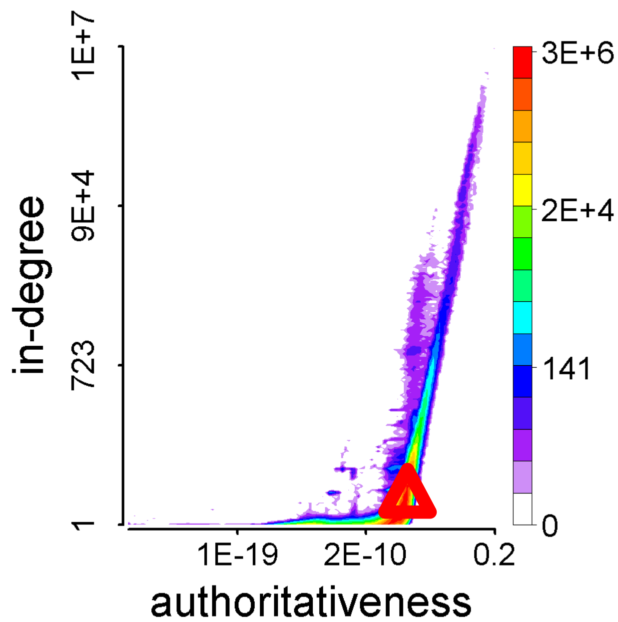


Jiang et al. **CatchSync: Catching Synchronized Behavior in Large Directed Graphs.** *KDD Best Paper Finalist*, 2014.

Synchronized Behavior: Normality

$$norm(u) = \frac{\sum_{(v,v') \in \mathcal{F}(u) \times \mathcal{U}} \mathbf{p}(v) \cdot \mathbf{p}(v')}{d(u) \times N}$$

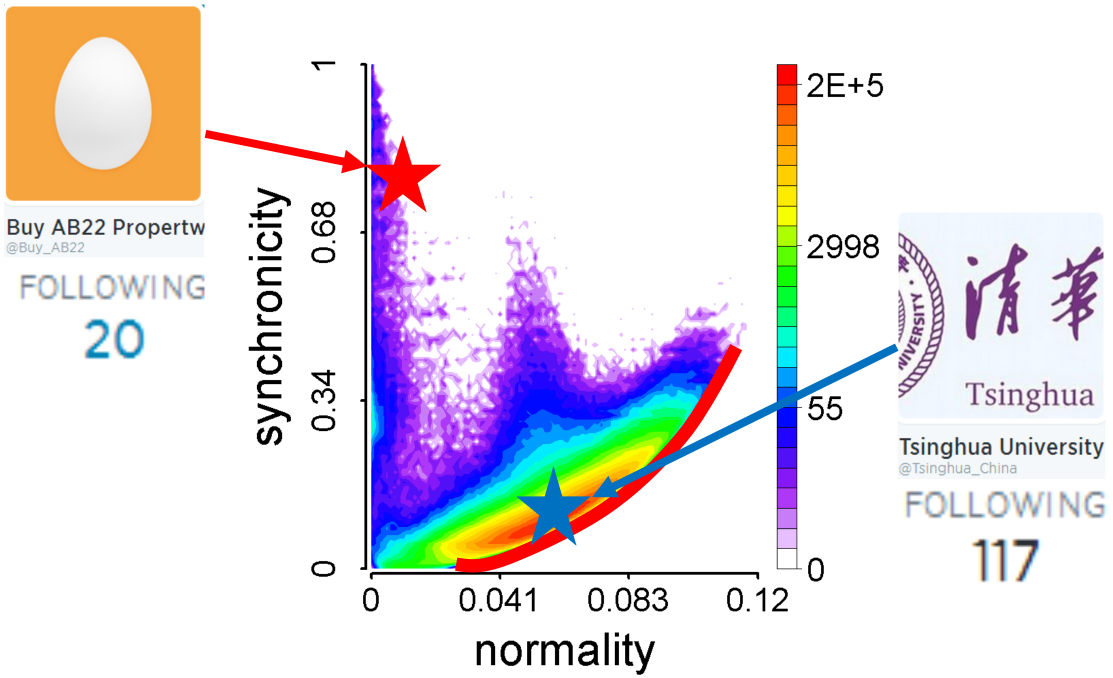




Synchronized Behavior: Theorem & CatchSync

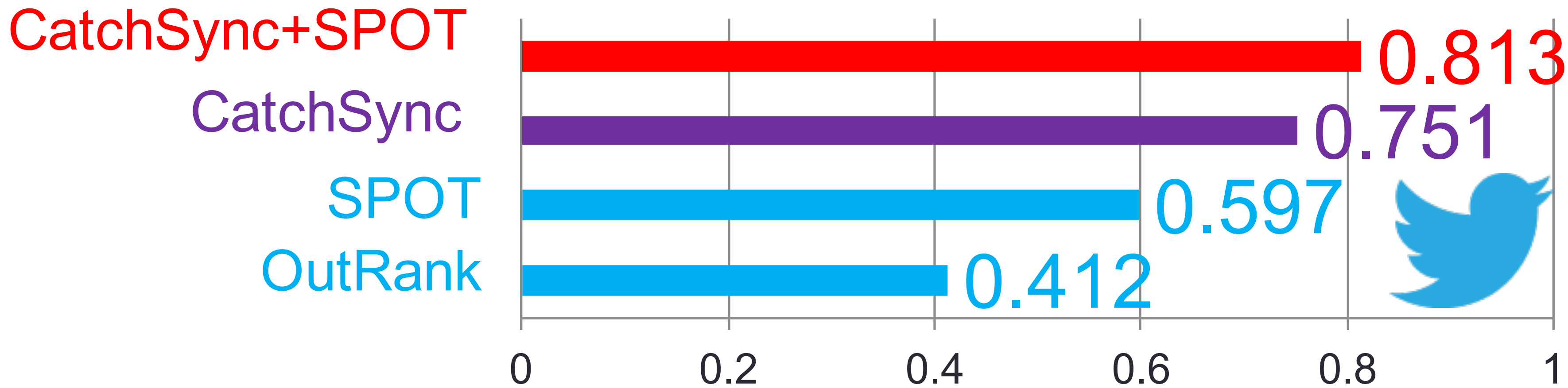
❖Synchronicity-Normality Plot

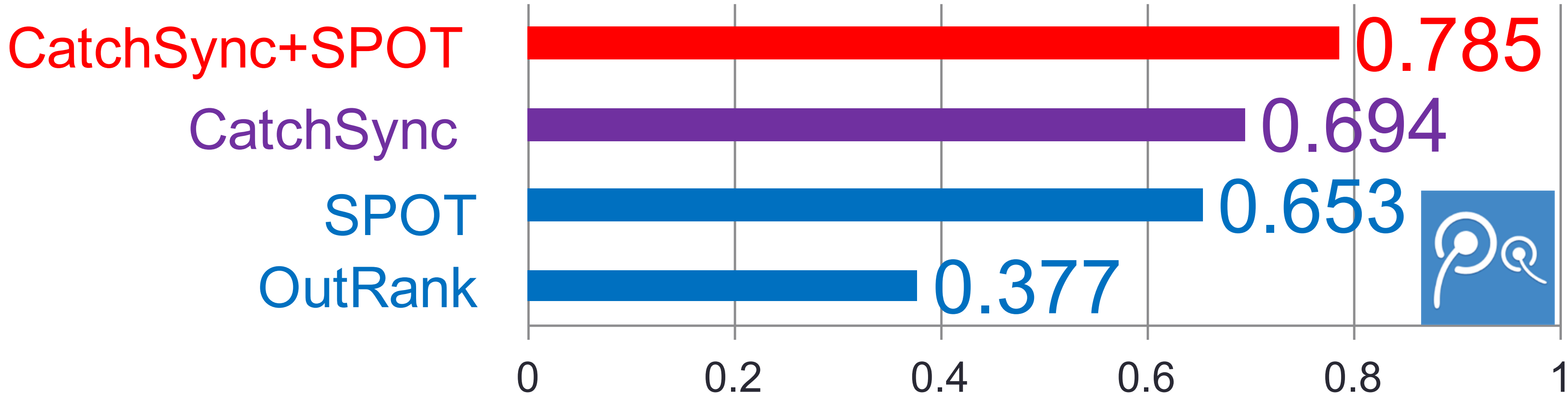
$$s_{min} = (-Mn^2 + 2n - s_b)/(1 - Ms_b)$$



Jiang et al. **CatchSync: Catching Synchronized Behavior in Large Directed Graphs.** *KDD Best Paper Finalist*, 2014.

Synchronized Behavior: Performance



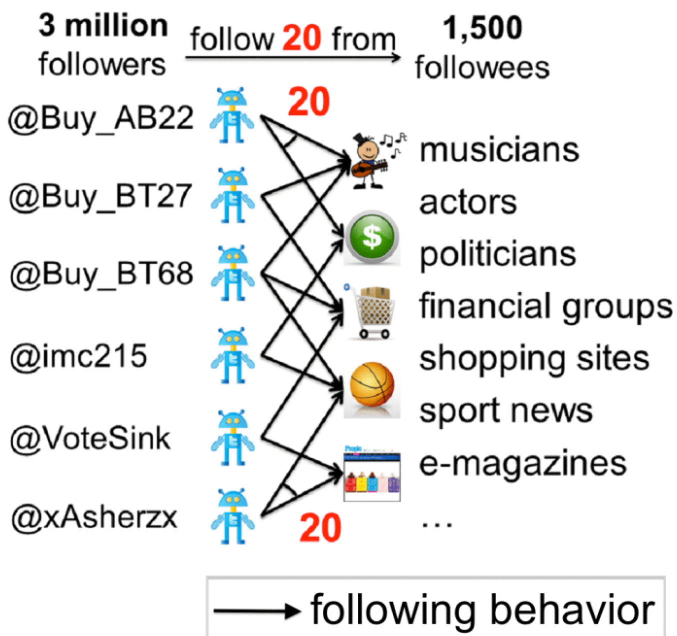


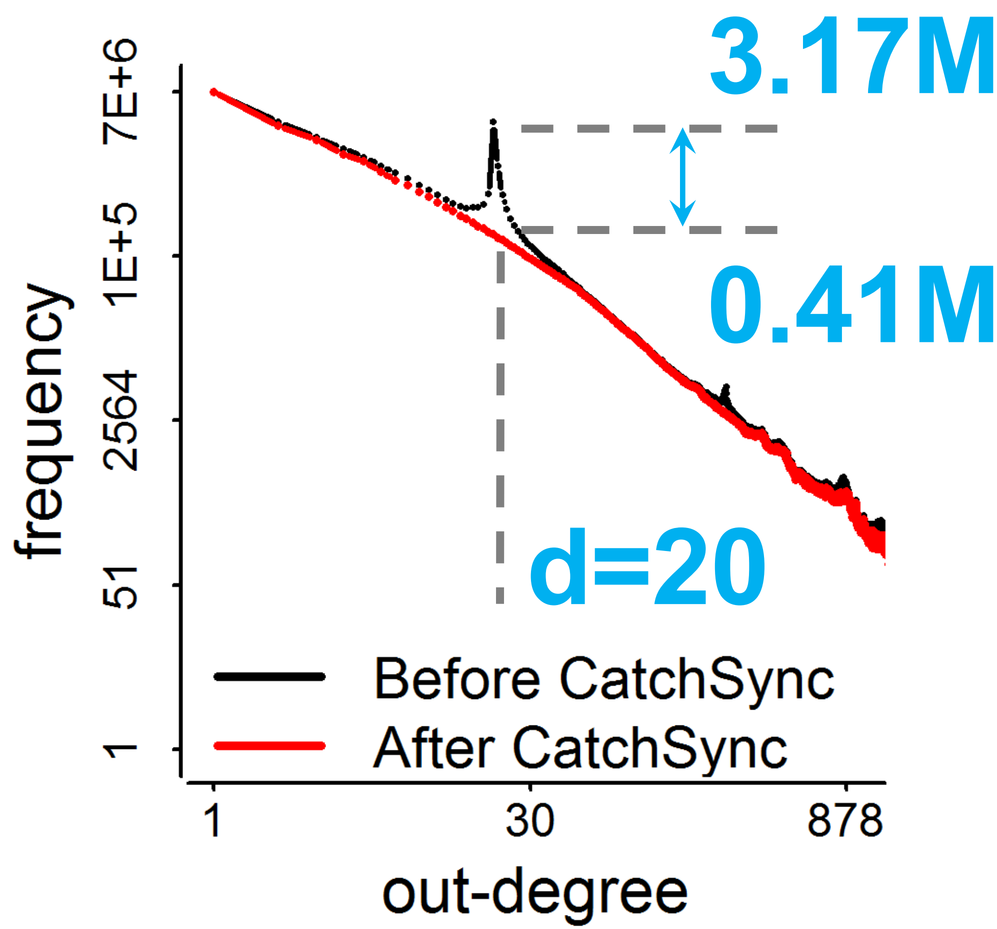
Jiang et al. **CatchSync: Catching Synchronized Behavior in Large Directed Graphs.** *KDD Best Paper Finalist*, 2014.

Synchronized Behavior: Performance

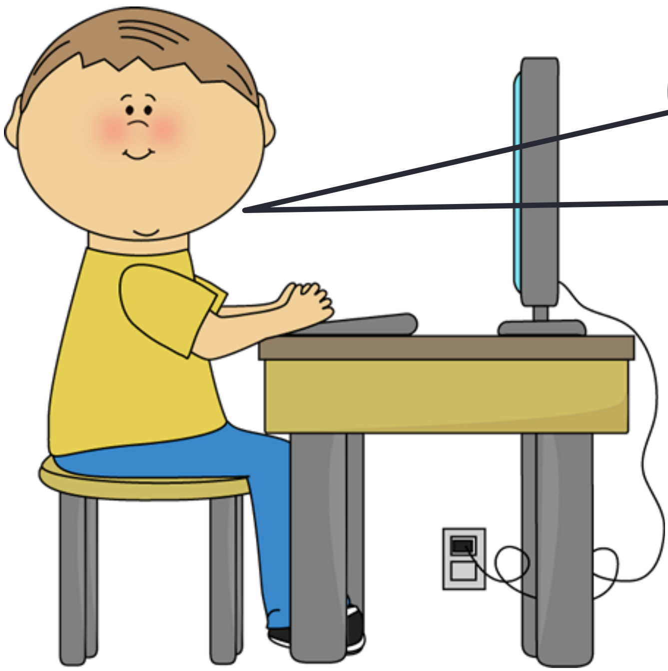
2009

41M





Beyond Graph: Multi-Dimensional Fraud



My boss wants me to **catch fraud** in such a big table – **billions of records, tens of columns**!!! How?!

	ID	USER_NAME	CREATED_AT	TEXT	HASH_TAGS
1	251	SpiritSofts	Dec 14, 2013	SAP HANA ONLINE TRAINING COURSE CONTENT http://t.co/2DefOMC0Vi	
2	252	Blue net studiO	Dec 14, 2013	sap hana online training and placenet 2 http://t.co/S1wGh8n5Kk	
3	253	Hana Kingham	Dec 14, 2013	Right film fest today: love actually, elf, gravity, training day. #dayyym	dayyym,
4	254	Nora Apnila J...	Dec 14, 2013	Alhamdulilaaahhhh...selesai ikutin kelanjutan training dadakan mb Hana ...	
5	255	ZaranTech	Dec 14, 2013	I added a video to a @YouTube playlist http://t.co/O3qD9wfi8K SAP BUSI...	
6	256	ZaranTech	Dec 14, 2013	I added a video to a @YouTube playlist http://t.co/XxrfuCUqAS SAP BUSI...	
7	257	Helmich op t...	Dec 14, 2013	Reserveer alvast 15 januari 2014 training HANA Essentials #SAP #HANA	SAP,HANA,
8	258	Social News	Dec 13, 2013	sap hana online training and placenet 2 http://t.co/JlaA41ldnV	
9	259	Nurjanah	Dec 13, 2013	Baca notif fb .. ada training dadakaann dari eyang kita.... ayo wara wiri ca...	
10	260	Nora Apnila J...	Dec 13, 2013	lanjutt di rumaah dulu ikutan trainingnyaaa..mau buru buru pulang see u...	
11	261	madhu	Dec 13, 2013	SAP HANA TRAINING | SAP HANA PLACEMENT | SAP HANA INSTITUTE I...	
12	262	Hana O'Neill	Dec 13, 2013	@sarahsilvanator no I have life guard training Saturday and my final test t...	
13	263	arjun	Dec 13, 2013	sap grc online training |sap hana sap security online training@YEKTEK - A...	

fraud

Jiang et al. **A General Suspicious Metric for Dense Blocks in Multi-Modal Data.** *ICDM*, 2015.

Multi-Dimensional Fraud

Wholesalebargain2015 Retweeted

Real Time Deals @ebayrt · 2h

Seattle Mariners Mlb #Majestic Authentic Diamond Blue Stitched 2 Xl M... (Sanford) USD 25 ebayrt.co/sports-mem-car… #eBay #eBayUS via @wil30225

2

Wholesalebargain2015 Retweeted

Real Time Deals @ebayrt · 2h

Embroidered Navy Blue Aztec Mexican Top/ Long Sleeve Ethnic Mod... USD 35 ebayrt.co/clothing-shoes… #Handmade #eBay #eBayUS via @smilingbluedog

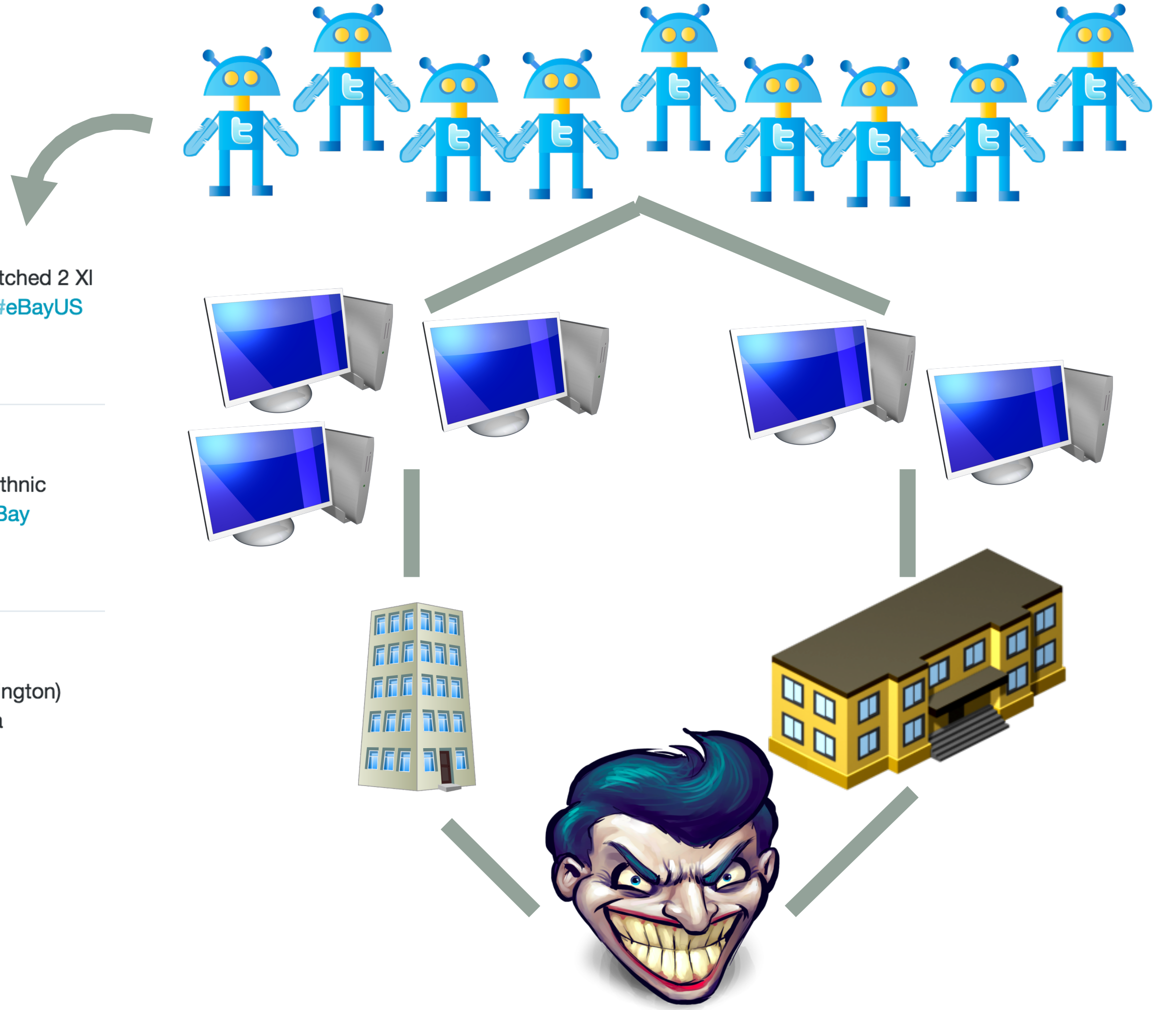
4

Wholesalebargain2015 Retweeted

Real Time Deals @ebayrt · 1h

Contractubex Children Cartoon Boxing Gloves Red (Bloomington) USD 21.78 ebayrt.co/sporting-goods… #eBay #eBayUS via @GaroldFrenz

4



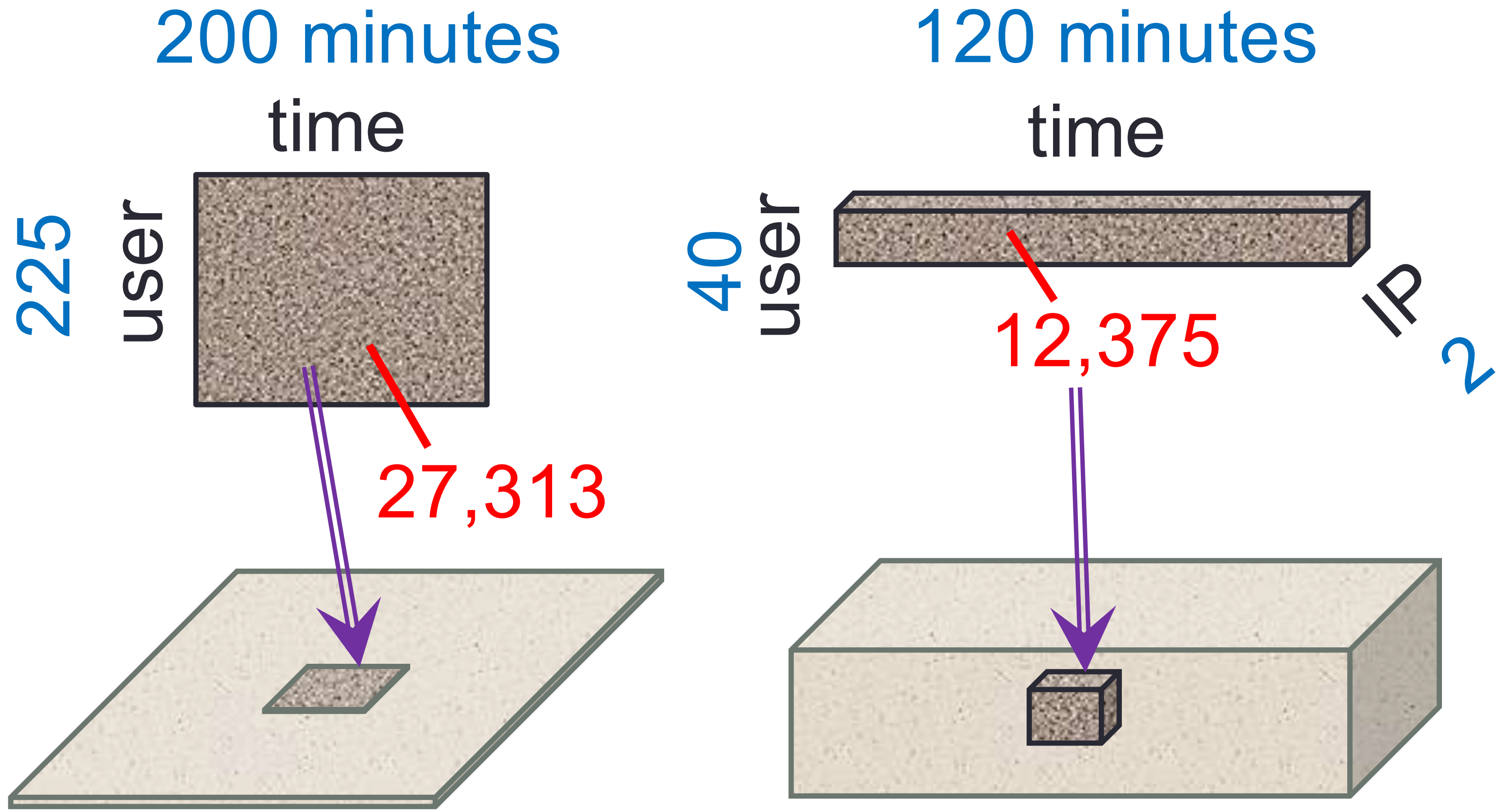
Jiang et al. **A General Suspicious Metric for Dense Blocks in Multi-Modal Data.** *ICDM*, 2015.

Multi-Dimensional Fraud

Dataset	Mode				Mass
Retweeting	User	Root ID	IP	Time (min)	#retweet
	29.5M	19.8M	27.8M	56.9K	211.7M
Trending (Hashtag)	User	Hashtag	IP	Time (min)	#tweet
	81.2M	1.6M	47.7M	56.9K	276.9M
Network attacks (LBNL)	Src-IP	Dest-IP	Port	Time (sec)	#packet
	2,345	2,355	6,055	3,610	230,836

Jiang et al. **A General Suspicious Metric for Dense Blocks in Multi-Modal Data.** *ICDM*, 2015.

Suspiciousness: Density in Multi-Dimensions

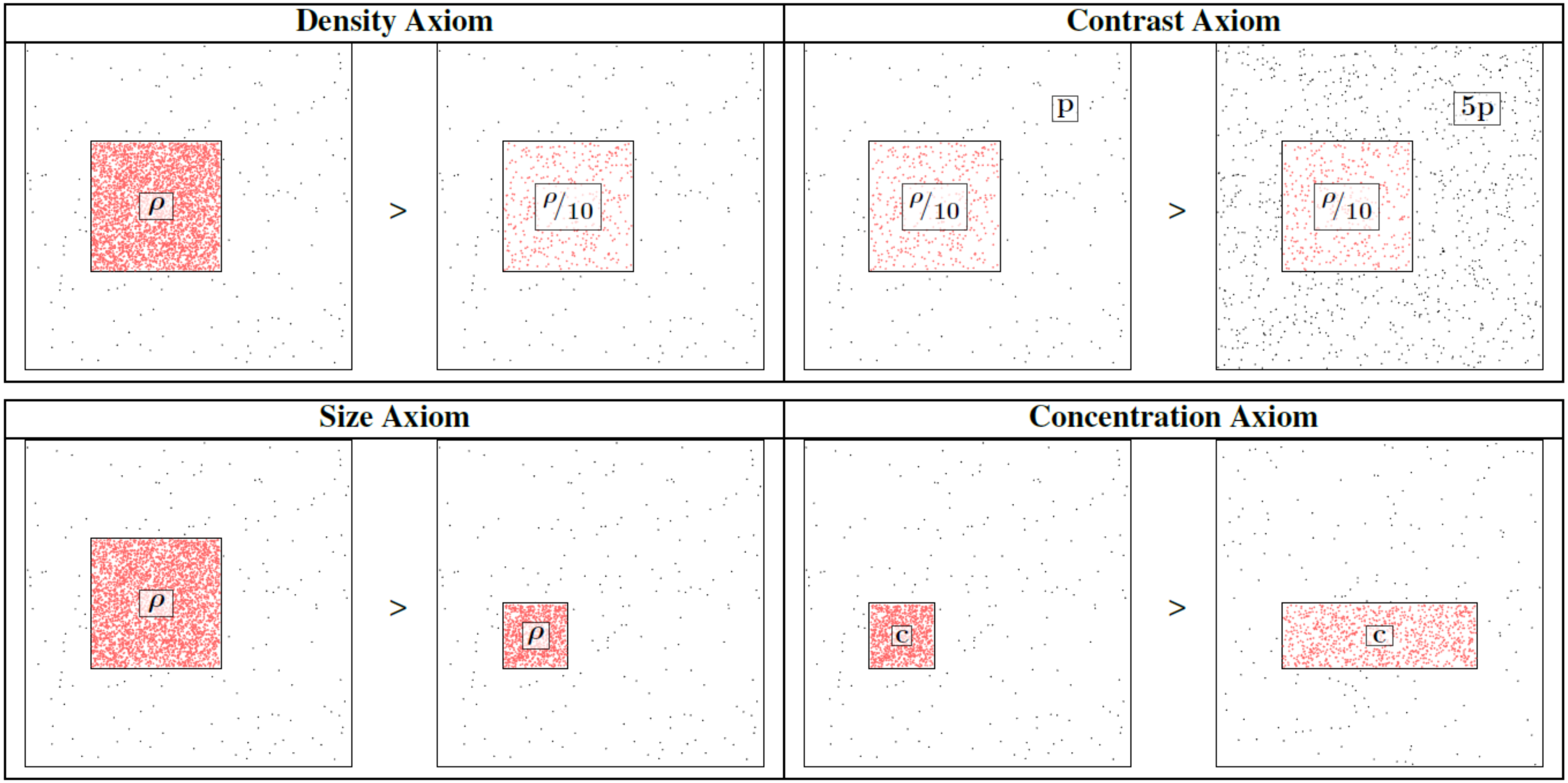


Question: Which is more suspicious?

- Before we search, we should be able to rank.

Jiang et al. **A General Suspicious Metric for Dense Blocks in Multi-Modal Data.** *ICDM*, 2015.

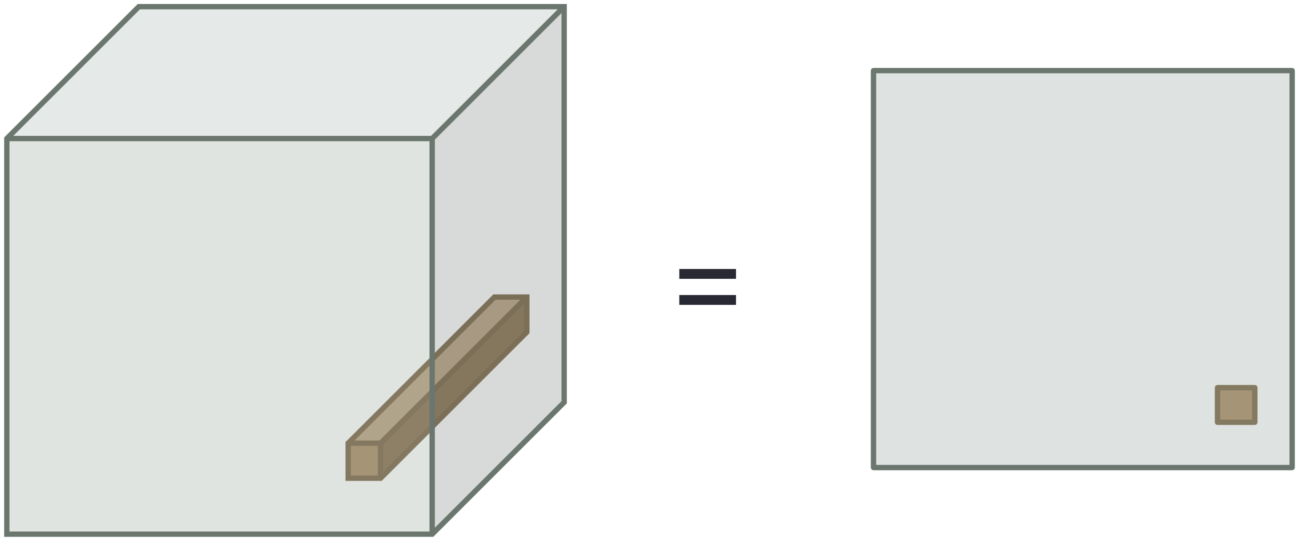
Suspiciousness: Axiom 1 - 4



Jiang et al. **A General Suspicious Metric for Dense Blocks in Multi-Modal Data.** *ICDM*, 2015.

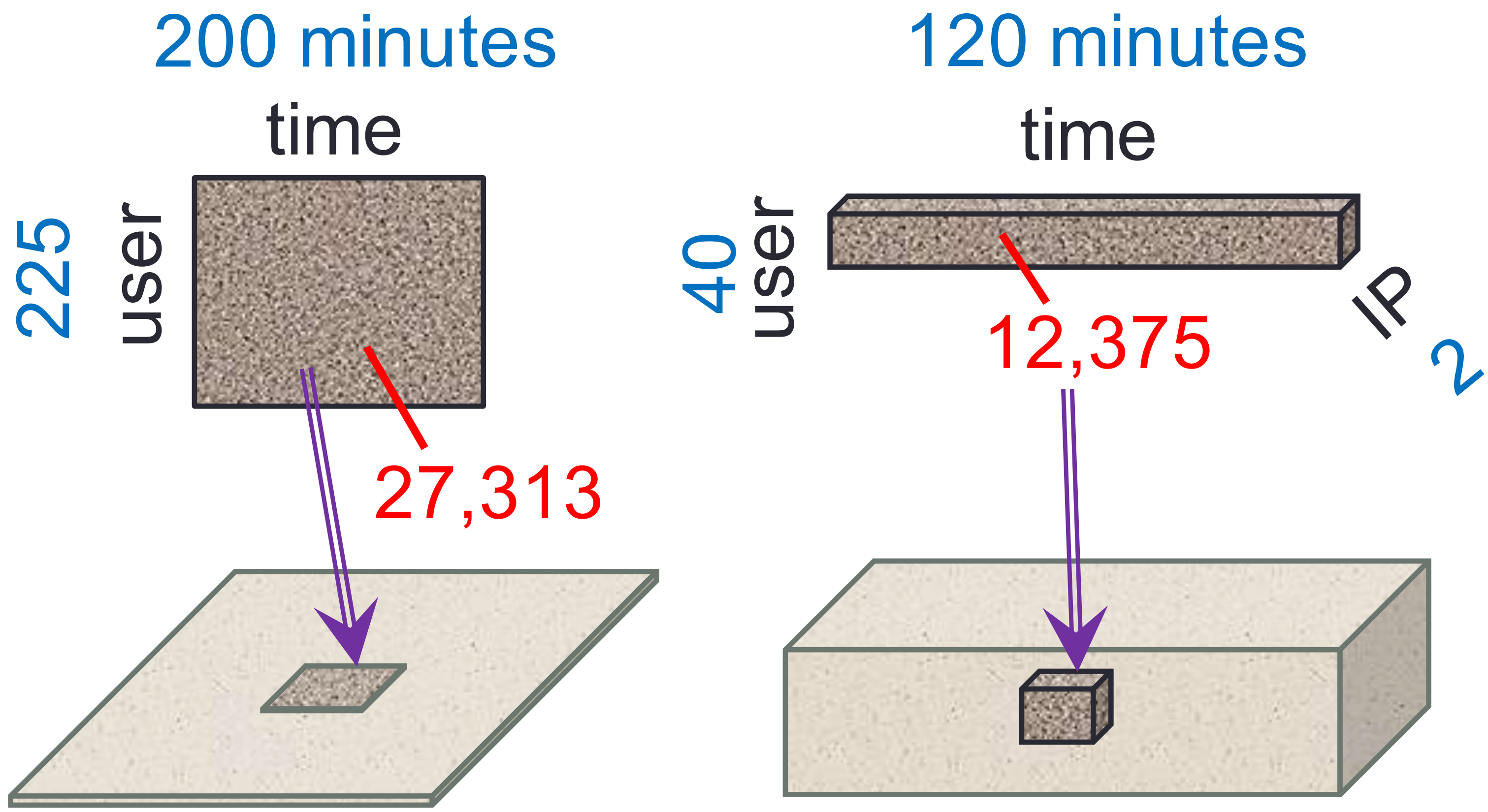
Suspiciousness: Axiom 5 Multimodal

Not including a mode is the same as including all values for that mode.



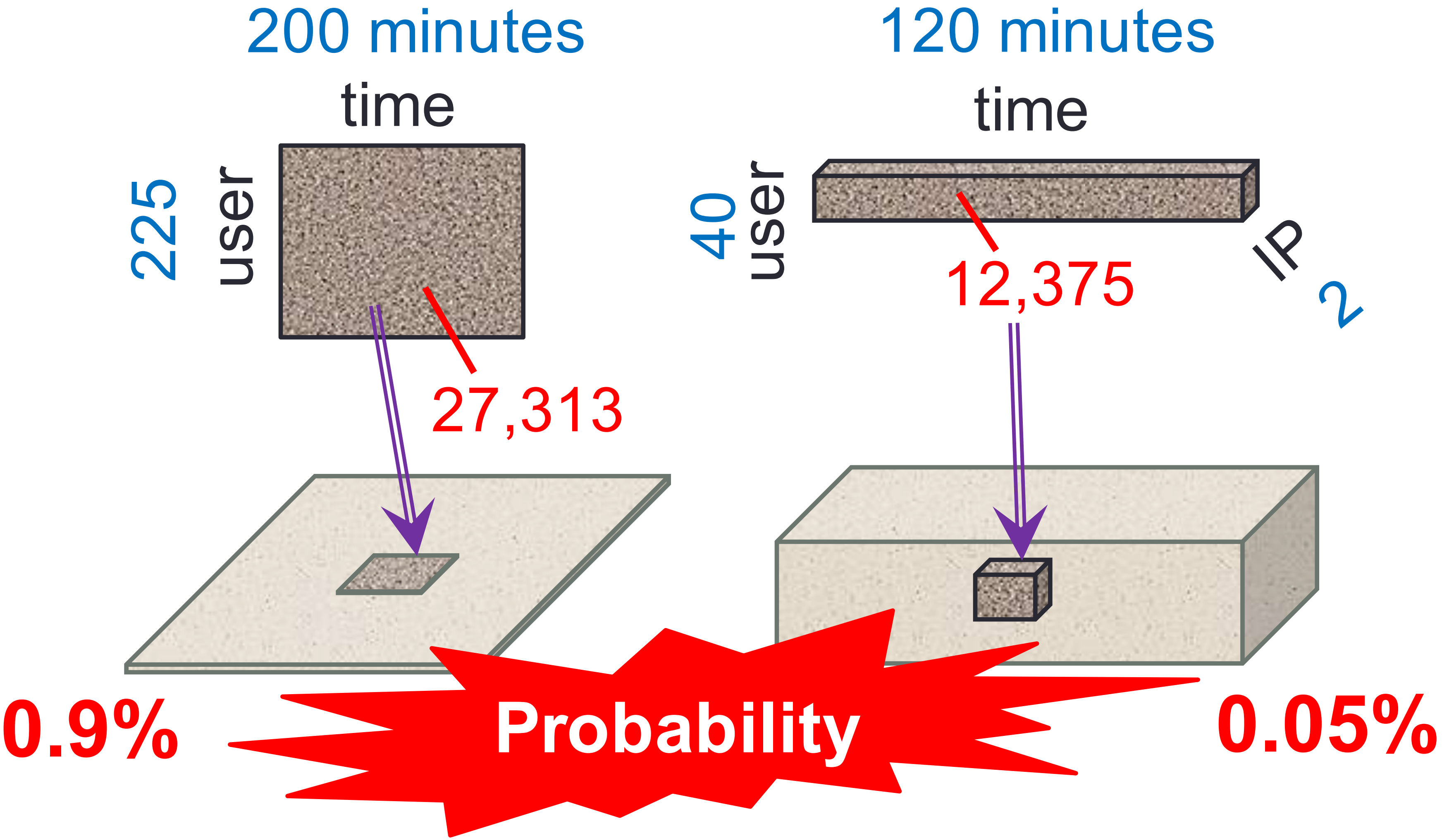
- New information (more modes) can only make our blocks more suspicious

Suspiciousness: General Metric & CrossSpot



Jiang et al. **A General Suspicious Metric for Dense Blocks in Multi-Modal Data.** *ICDM*, 2015.

Suspiciousness: General Metric & CrossSpot



Jiang et al. **A General Suspicious Metric for Dense Blocks in Multi-Modal Data.** *ICDM*, 2015.

Suspiciousness: General Metric & CrossSpot

- Negative log likelihood of block's probability

$$f(n, c, N, C) = -\log\left[Pr(Y_n = c)\right]$$

Lemma *Given an $n_1 \times \cdots \times n_K$ block of mass c in $N_1 \times \cdots \times N_K$ data of total mass C, the suspiciousness function is*

$$\boxed{f(\mathbf{n}, c, \mathbf{N}, C) = c(\log\frac{c}{C} - 1) + C\prod_{i=1}^{K}\frac{n_i}{N_i} - c\sum_{i=1}^{K}\log\frac{n_i}{N_i}}$$

Using ρ as the block's density and p is the data's density, we have the simpler formulation

$$\boxed{\hat{f}(\mathbf{n}, \rho, \mathbf{N}, p) = \left(\prod_{i=1}^{K} n_i\right) D_{KL}(\rho||p)}$$

Suspiciousness: Trend Manipulating

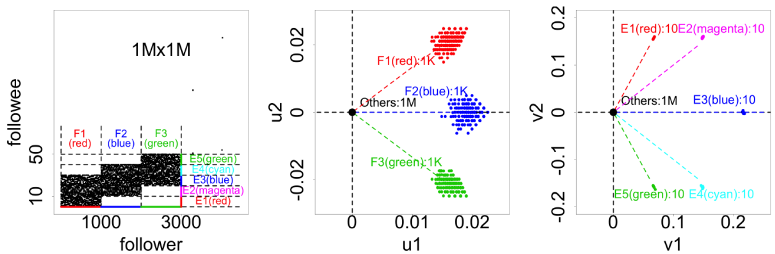
User×hashtag×IP×minute	Mass c	Suspiciousness
582×3×294×**56,940**	5,941,821	111,799,948
188×1×313×**56,943**	2,344,614	47,013,868
75×1×2×2,061	689,179	19,378,403

User ID	Time	IP address (city, province)	Tweet text with hashtag
USER-D	11-18 12:12:51	IP-1 (Deyang, Shandong)	**#Snow#** the Samsung GALAXY SII QQ Service customized version...
USER-E	11-18 12:12:53	IP-1 (Deyang, Shandong)	**#Snow#** the Samsung GALAXY SII QQ Service customized version...
USER-F	11-18 12:12:54	IP-2 (Zaozhuang, Shandong)	**#Snow#** the Samsung GALAXY SII QQ Service customized version...
USER-E	11-18 12:17:55	IP-1 (Deyang, Shandong)	**#Li Ning - a weapon with a hero#** good support activities!
USER-F	11-18 12:17:56	IP-2 (Zaozhuang, Shandong)	**#Li Ning - a weapon with a hero#** good support activities!
USER-D	11-18 12:18:40	IP-1 (Deyang, Shandong)	**#Toshiba Bright Daren#** color personality test to find out your sense...
USER-E	11-18 17:00:31	IP-2 (Zaozhuang, Shandong)	**#Snow#** the Samsung GALAXY SII QQ Service customized version...
USER-D	11-18 17:00:49	IP-2 (Zaozhuang, Shandong)	**#Toshiba Bright Daren#** color personality test to find out your sense...
USER-F	11-18 17:00:56	IP-2 (Zaozhuang, Shandong)	**#Li Ning - a weapon with a hero#** good support activities!

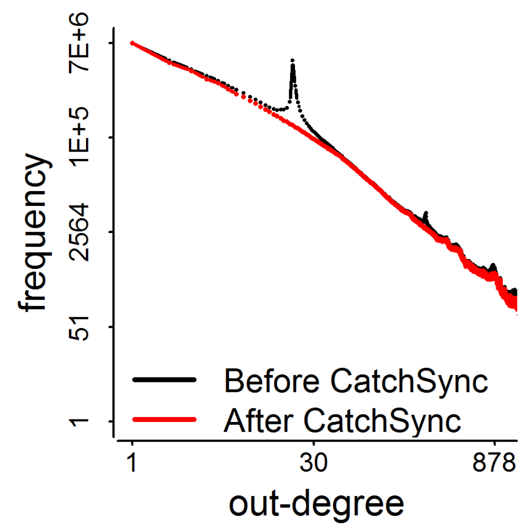
Jiang et al. **A General Suspicious Metric for Dense Blocks in Multi-Modal Data.** *ICDM*, 2015.

Summary for Detecting Unnatural Behavior

- Suspicious behavioral patterns
  - **Lockstep pattern:** CopyCatch, LockInfer
  - **Synchronized pattern:** CatchSync
  - **Multi-dimensional suspiciousness:** CrossSpot



Rule 4 ("pearls"): a "staircase" of three partially overlapping blocks.



TAKE AWAY

Behavior Modeling in Social Networks

Six Disruptive Basic Research Areas, by DOD

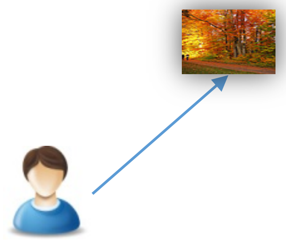
Basic Research Areas

- Six Disruptive Basic Research Areas
  - Engineered Materials (metamaterials and plasmonics)
  - Quantum Information and Control
  - Cognitive Neuroscience
  - Nanoscience and Nanoengineering
  - Synthetic Biology
  - Computational Modeling of Human and Social Behavior

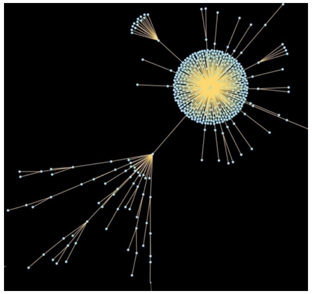
From Micro to Macro

Natural behavior

Individual behavior



Information cascade



Unnatural/Suspicious behavior

Modeling Individual Behavior

- ❖Pattern discovery, prediction and social recommendation
- ❖ContextMF: Social contexts (preference & influence)
- ❖FEMA: Spatiotemporal contexts (multi-faceted & dynamic)
- ❖Hybrid Random Walk: Social bridging multiple domains

Modeling Cascading Behavior

- Before predicting information spreading, understanding the behavioral mechanism is critical and fundamental.
- Behaviors can be modeled in different granularities, which depends on the target problem.
- Modeling information spreading with continuous-time model is promising and demonstrated to be effective in our research.

Detecting Unnatural Behavior

- Suspicious behavioral patterns
  - Lockstep pattern: CopyCatch, LockInfer
  - Synchronized pattern: CatchSync
  - Multi-dimensional suspiciousness: CrossSpot

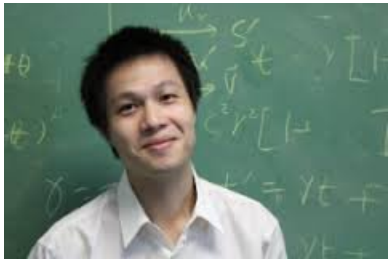
Acknowledge

Prof. Shiqiang Yang
Tsinghua

Prof. Fei Wang
University of Connecticut

Prof. Christos Faloutsos
CMU

Prof. Chaoming Song
Miami University



Prof. Wenwu Zhu
Tsinghua

Shifei Jin
Tsinghua

Alex Beutel
CMU

Bryan Hooi
CMU

THANK YOU!

Behavior Modeling in Social Networks