

MINING BEHAVIORAL NETWORKS

MODELING COMPLEX BEHAVIORS
IN SOCIAL MEDIA

Meng Jiang

mjiang89@gmail.com

www.meng-jiang.com

ROADMAP

Background

Previous works (6 problems)

Thinking & Future works

Behavior-based Applications/Systems

- Values, benefits and profits in social media



Post, forward text/image

Update Status | Add Photos/Video | Create Photo Album

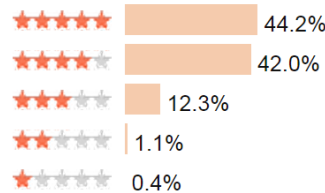
What's on your mind?

Public | Post

News feed ranking



Give ratings to movies



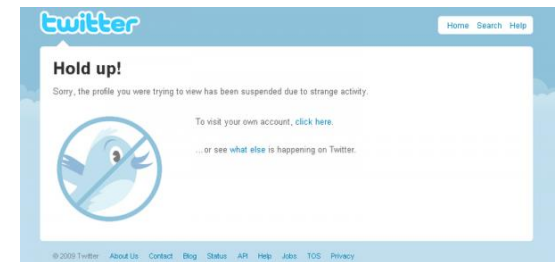
Recommender systems



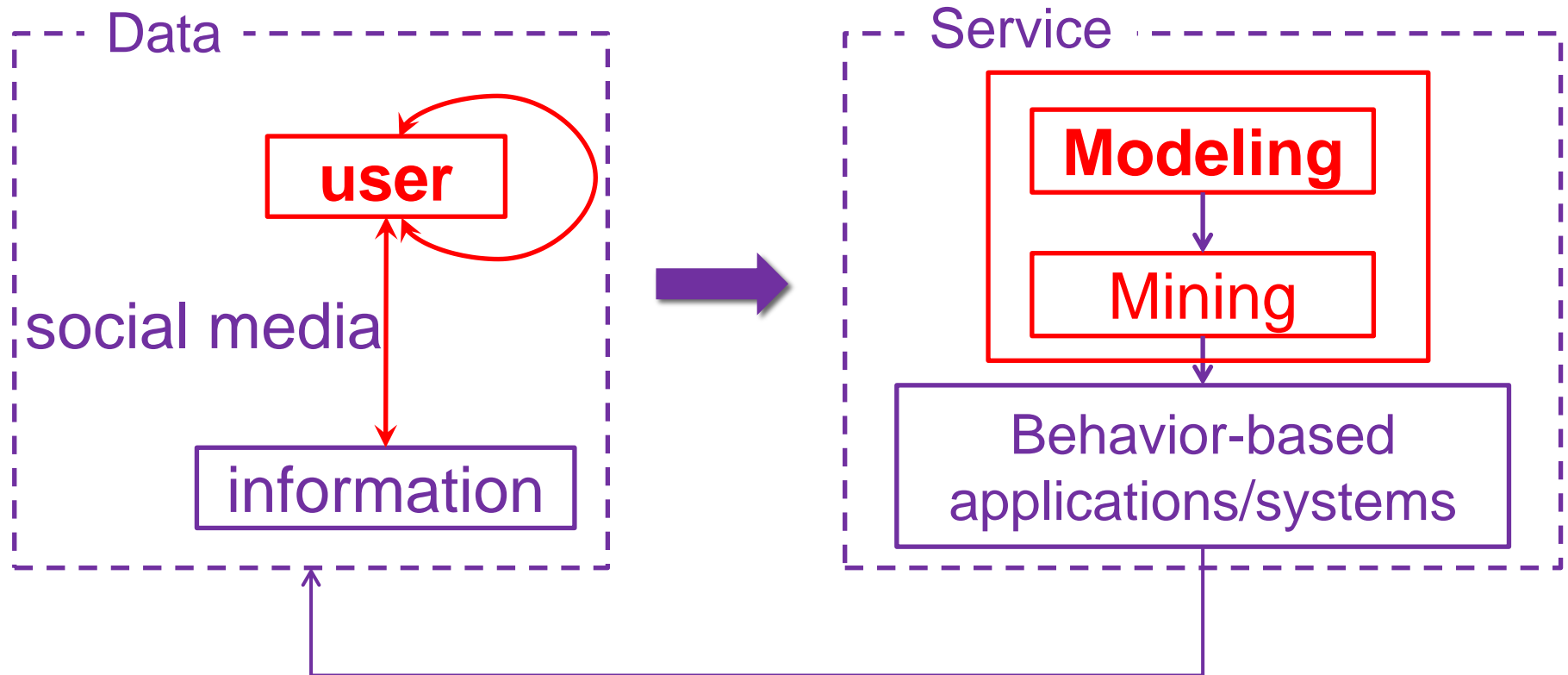
Zombie followers, fraud



Anti-spam, anti-fraud



Behavior Modeling in Social Media



ROADMAP

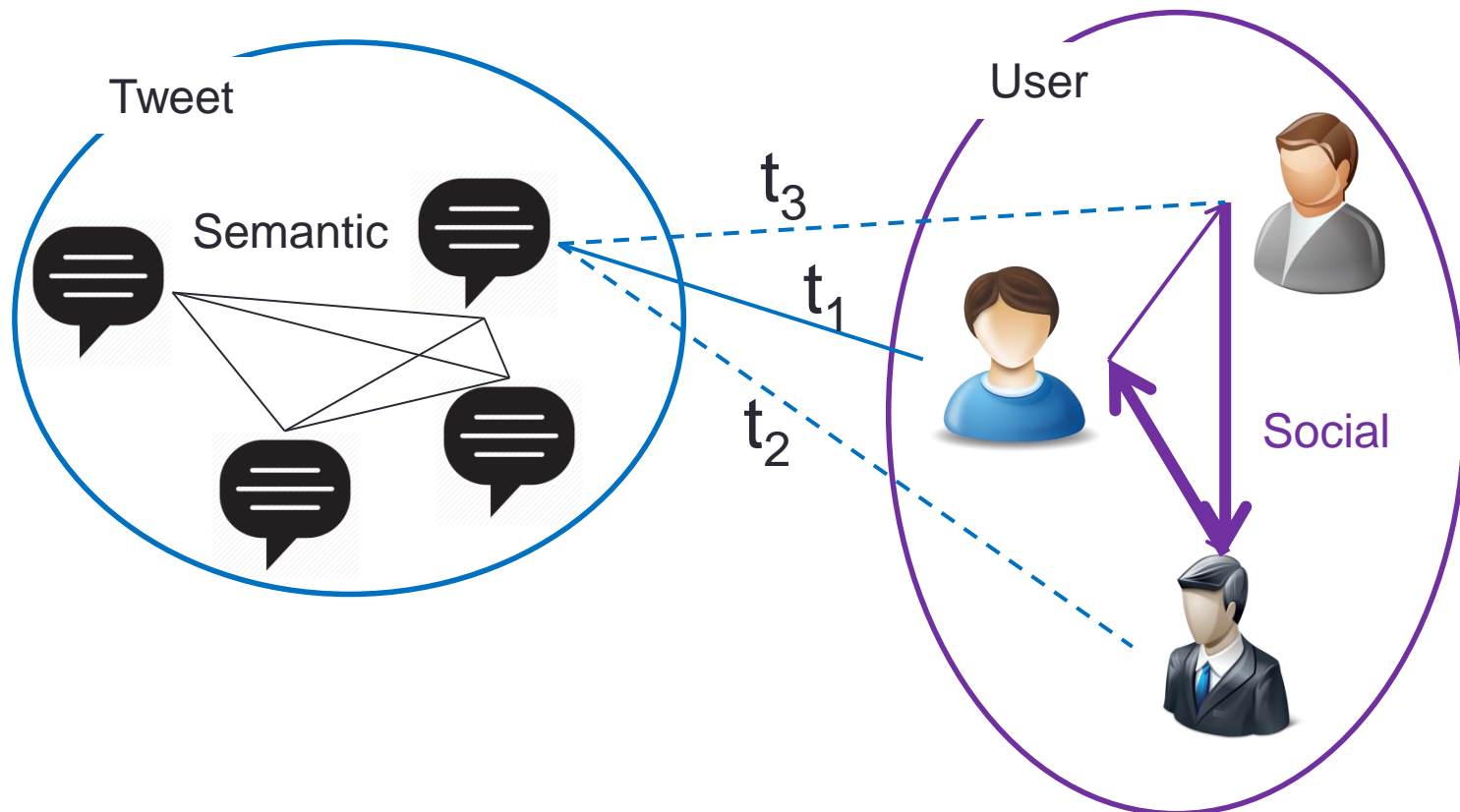
Background

Previous works (6 problems)

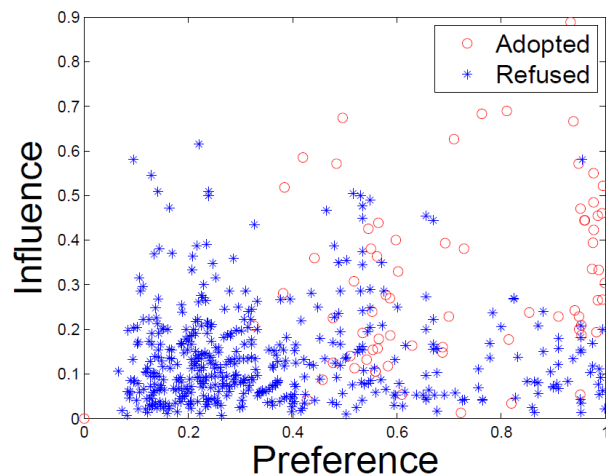
Thinking & Future works

FROM Behavior Modeling TO Mining Behavioral Networks

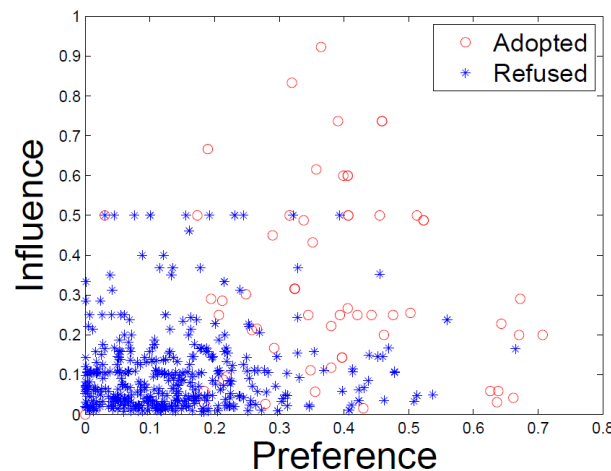
- Problem 1. **Information adoption behavior prediction**
(predicting who will retweet/share which message)
[CIKM'12a, TKDE'14]



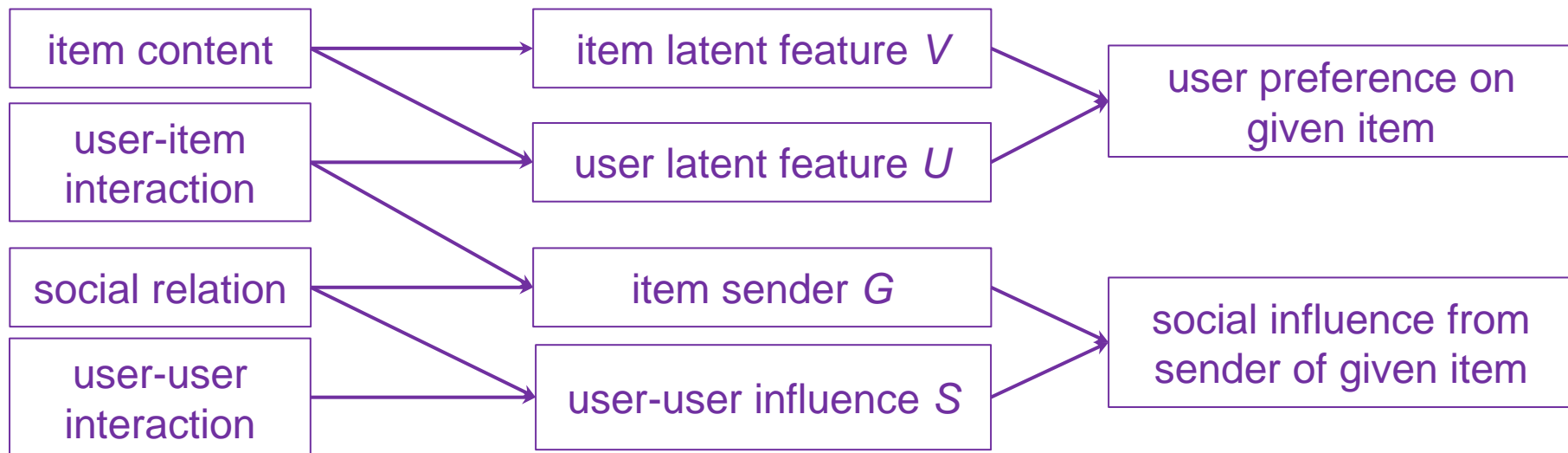
Social Contextual Behavioral Pattern



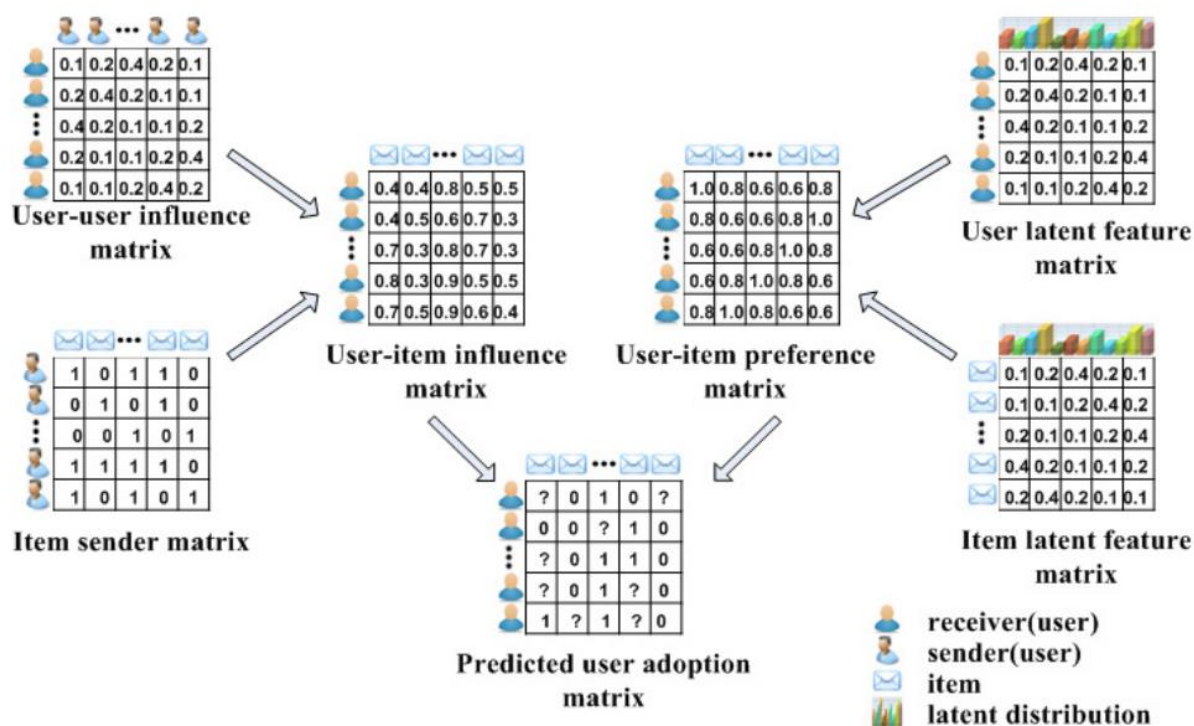
Renren



Tencent Weibo



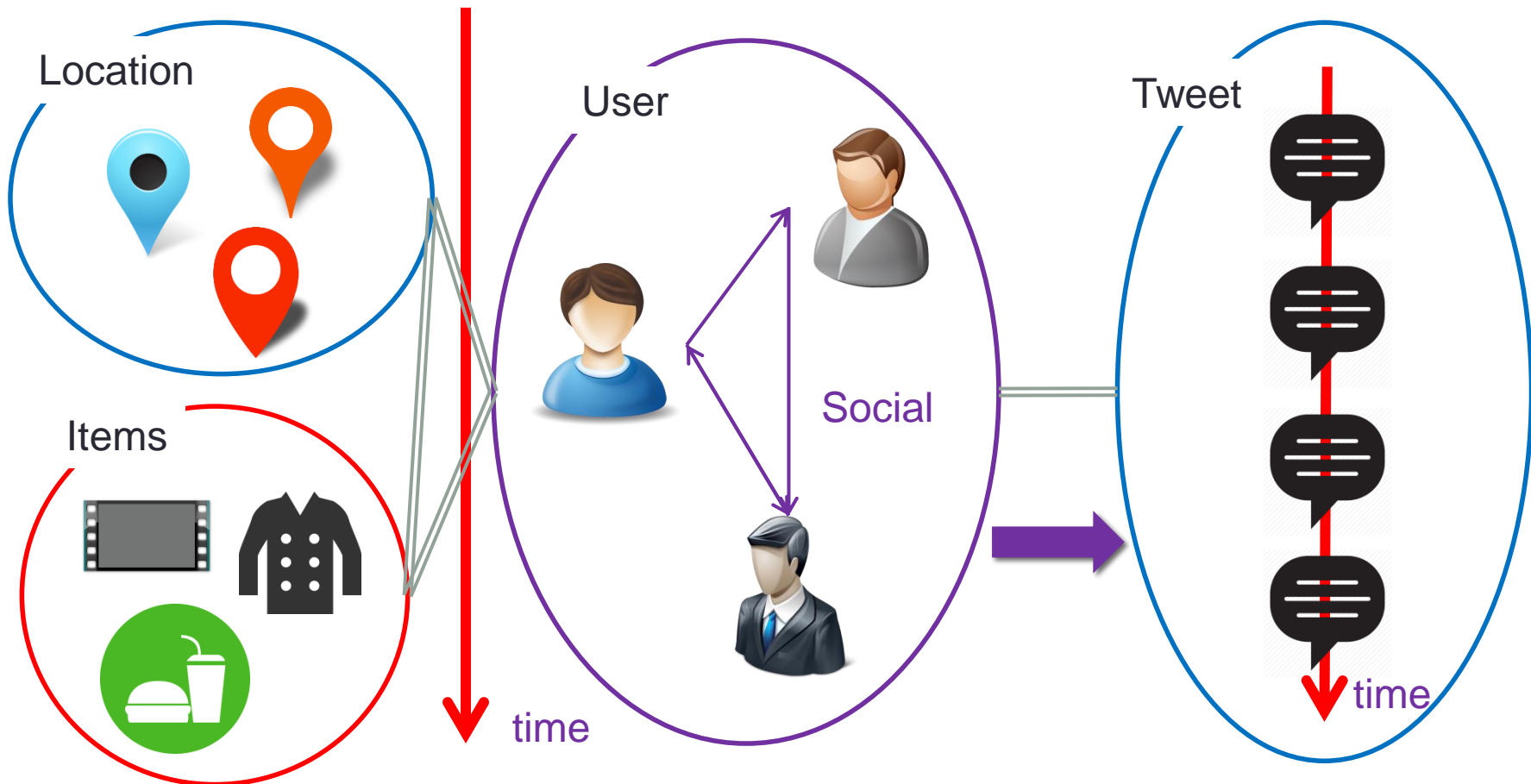
ContextMF: Social Contextual Model



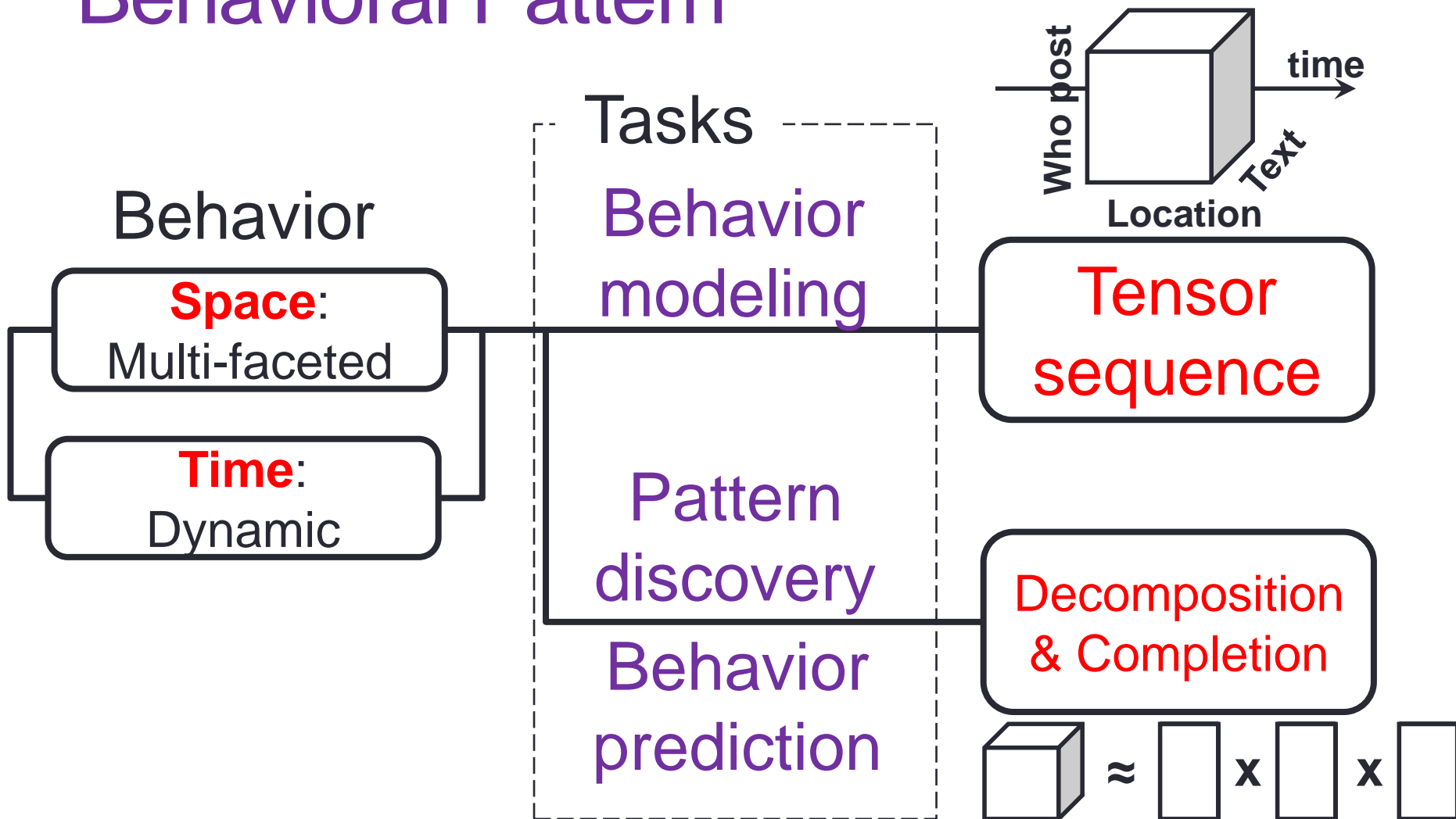
Method	MAE	RMSE
Renren Dataset		
Content-based [1]	0.3842	0.4769
Item CF [25]	0.3601	0.4513
FeedbackTrust [22]	0.3764	0.4684
Influence-based [9]	0.3859	0.4686
SoRec [19]	0.3276	0.4127
SoRec [20]	0.2985	0.3537
Influence MF	0.3102	0.3771
Preference MF	0.3032	0.3762
Context MF	0.2416	0.3086
Tencent Weibo Dataset		
Content-based [1]	0.2576	0.3643
Item CF [25]	0.2375	0.3372
FeedbackTrust [22]	0.2830	0.3887
Influence-based [9]	0.2651	0.3813
SoRec [19]	0.2256	0.3325
SoRec [20]	0.1997	0.2962
Influence MF	0.2183	0.3206
Preference MF	0.2111	0.3088
Context MF	0.1514	0.2348

FROM Behavior Modeling TO Mining Behavioral Networks

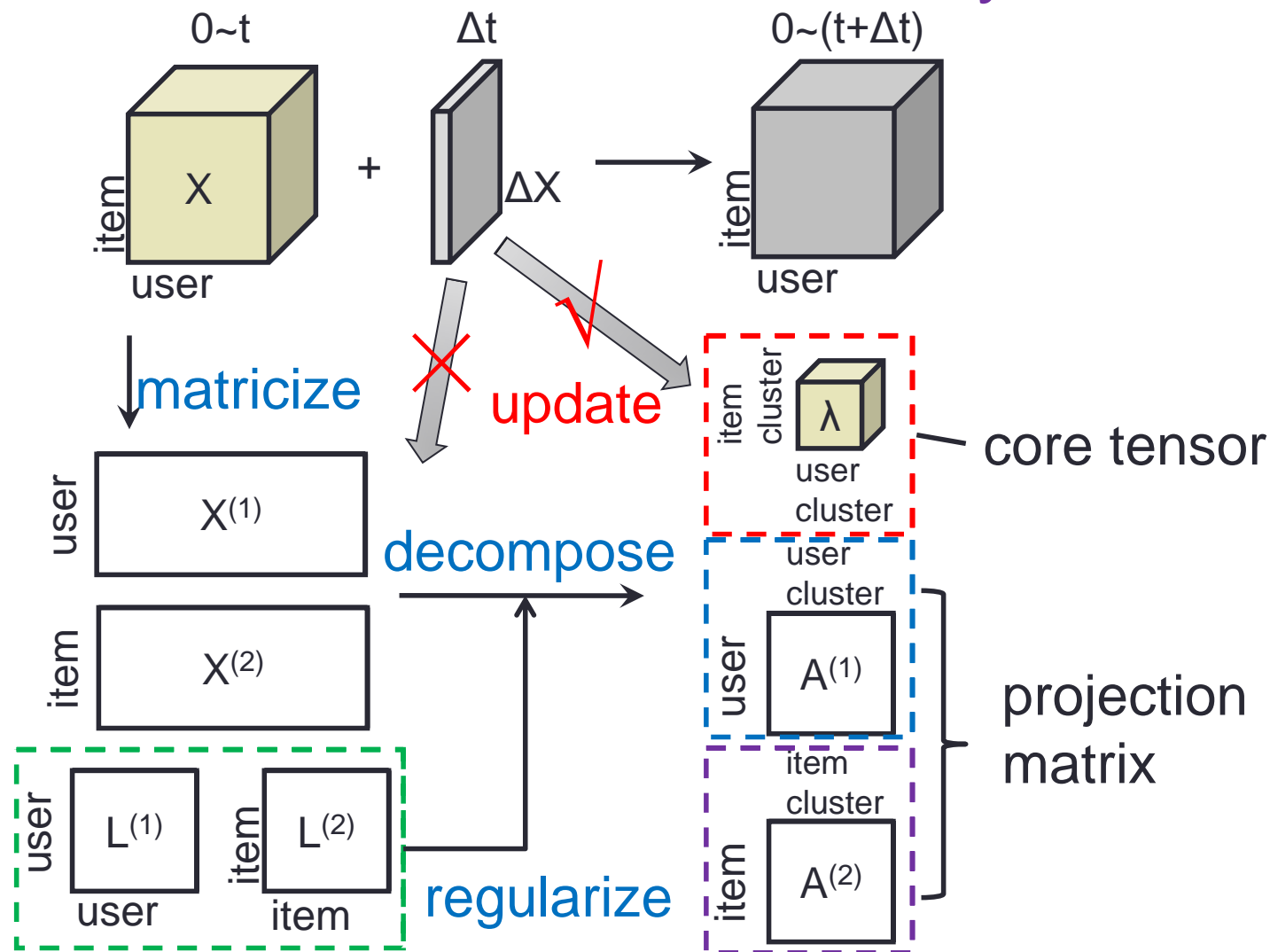
- Problem 2. **Content generation behavior prediction**
(predicting who will generate what content) [KDD'14]



Spatial Temporal Contextual Behavioral Pattern

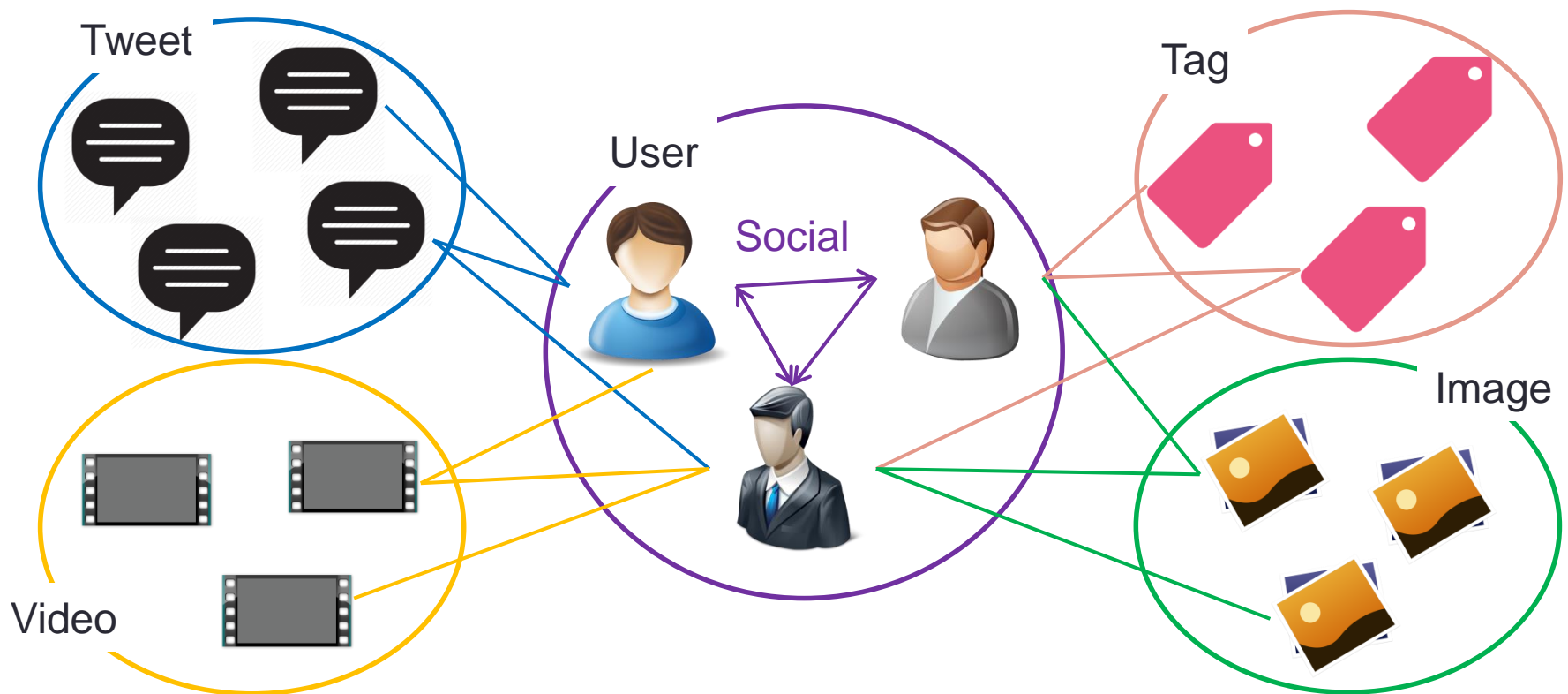


FEMA: Flexible Evolutionary Multi-faceted Analysis on Tensor Perturbation Theory



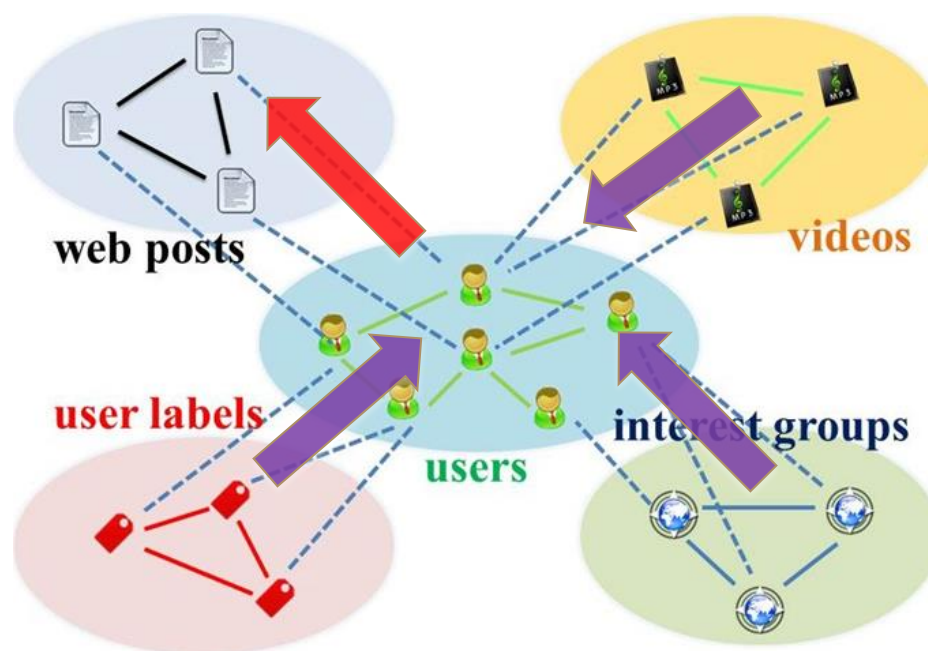
FROM Behavior Modeling TO Mining Behavioral Networks

- Problem 3. **Cross-domain behavior prediction**
(addressing cold-start problems if users adopt multiple types of items) [CIKM'12b, TKDE'15]



Hybrid Random Walk

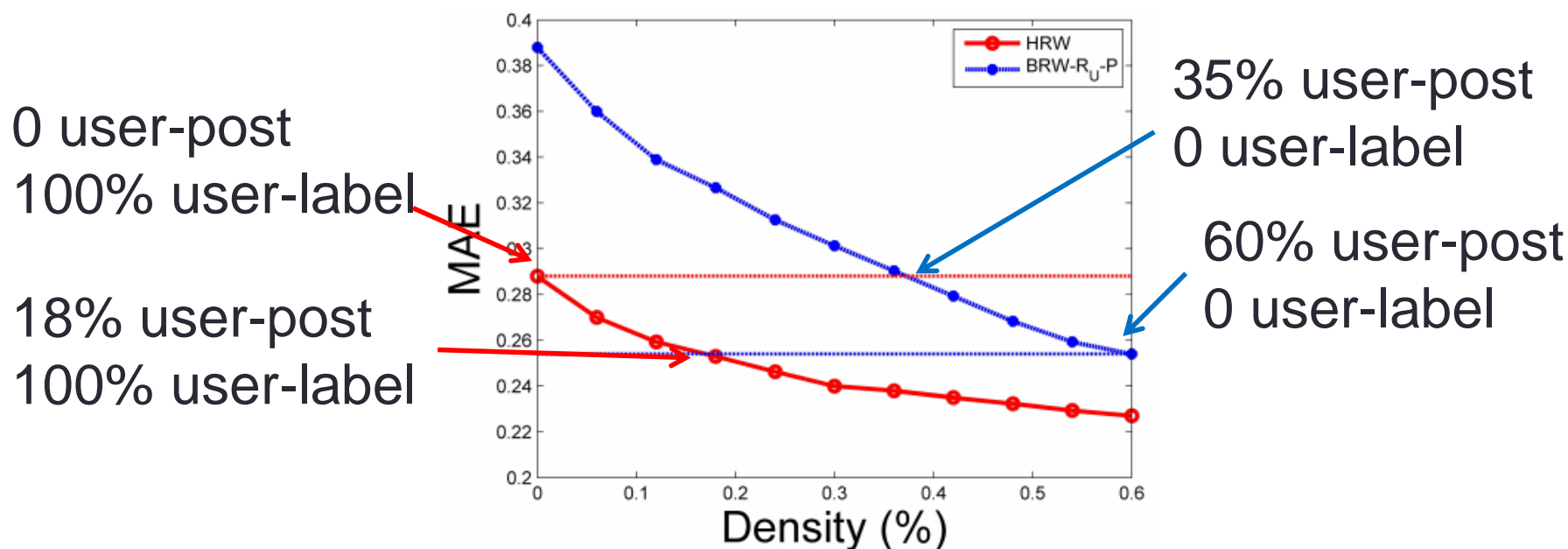
- Star-structured graph with social domain in the center
- Auxiliary domain \rightarrow Social domain \rightarrow Target domain



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

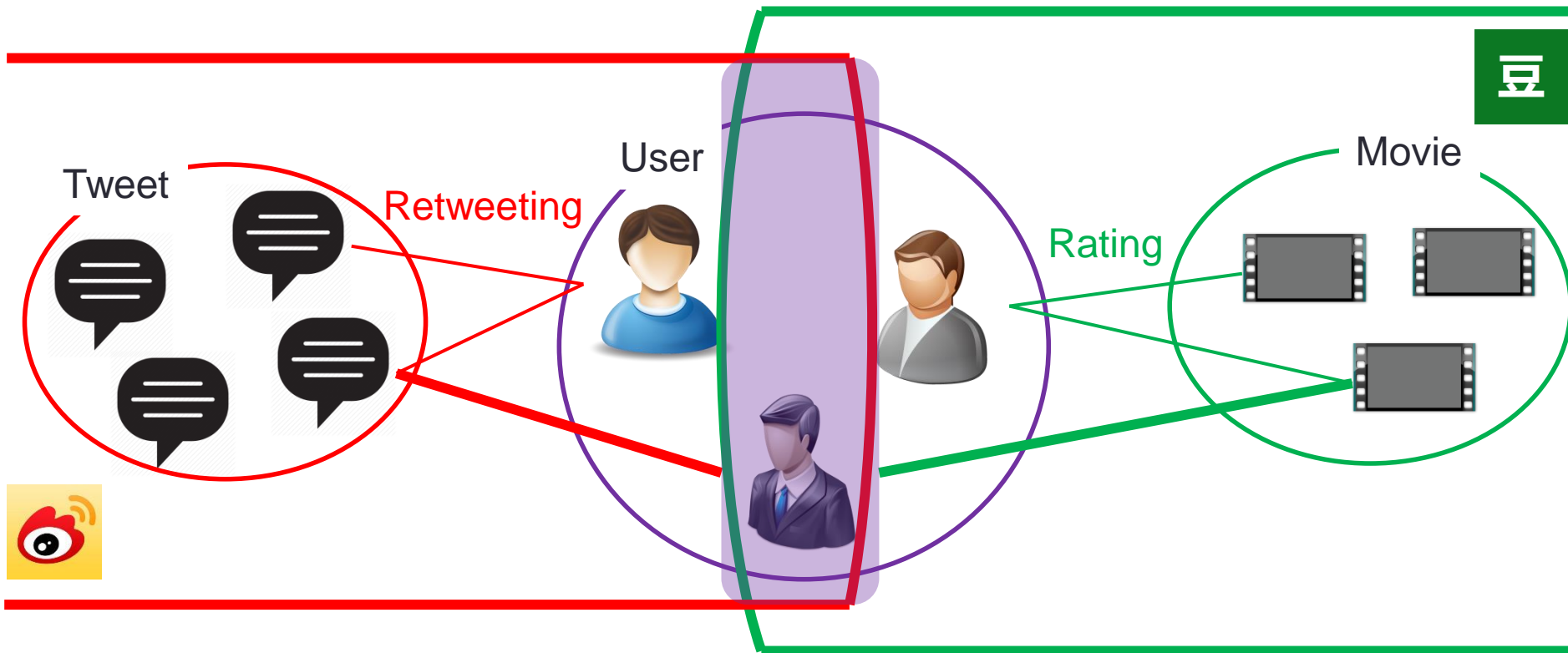
Performance on Cold-start Users

- Performance(“social label” data + 3-day “web post” data)
= Performance(10-day “web post” data)



FROM Behavior Modeling TO Mining Behavioral Networks

- Problem 4. **Cross-platform behavior prediction**
(addressing cold-start problems if platform A and platform B have overlapped users) [New]



Cross-platform Behavior Modeling with Semi-Supervised MF

■ Input

- Tgt./Aux. platform $\underline{P}/\underline{Q}$;
- Behavior data $\underline{R}^{(P)}/\underline{R}^{(Q)}$;
- Observation $\underline{W}^{(P)}/\underline{W}^{(Q)}$;
- Overlapping indicator $\underline{W}^{(P,Q)}$,

■ Output

- User latent representation $\underline{U}^{(P)}/\underline{U}^{(Q)}$;
- Item latent representation $\underline{V}^{(P)}/\underline{V}^{(Q)}$;
- Missing values in $\underline{R}^{(P)}$

■ Objective function

Target platform

Auxiliary platform

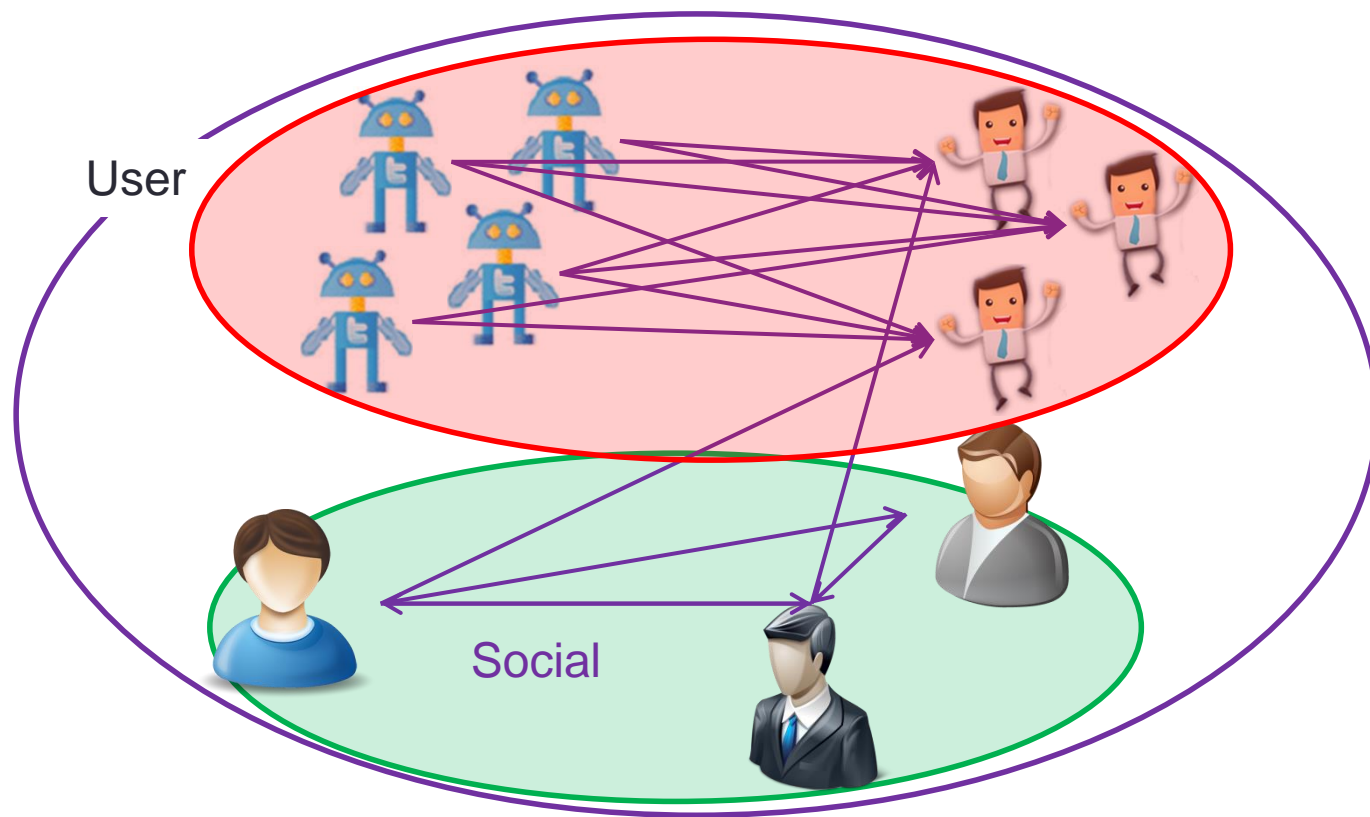
$$\begin{aligned}\mathcal{J} = & \sum_{i,j} W_{i,j}^{(P)} \left(R_{i,j}^{(P)} - \sum_r U_{i,r}^{(P)} V_{r,j}^{(P)} \right)^2 \\ & + \lambda \sum_{i,j} W_{i,j}^{(Q)} \left(R_{i,j}^{(Q)} - \sum_r U_{i,r}^{(Q)} V_{r,j}^{(Q)} \right)^2 \\ & + \mu \sum_{i_1,j_1,i_2,j_2} W_{i_1,j_1}^{(P,Q)} W_{i_2,j_2}^{(P,Q)}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

Pair-wise similarity
of overlapped users

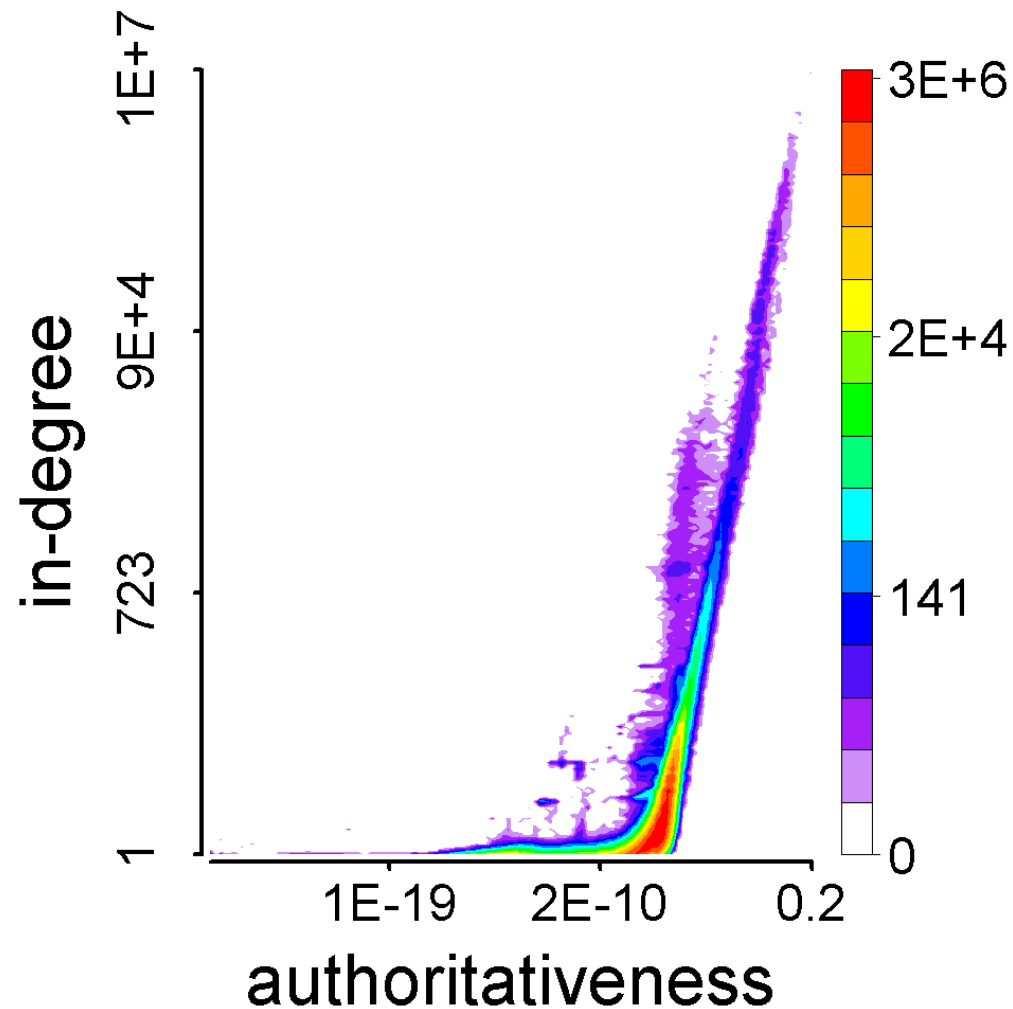
Supervised term

FROM Behavior Modeling TO Mining Behavioral Networks

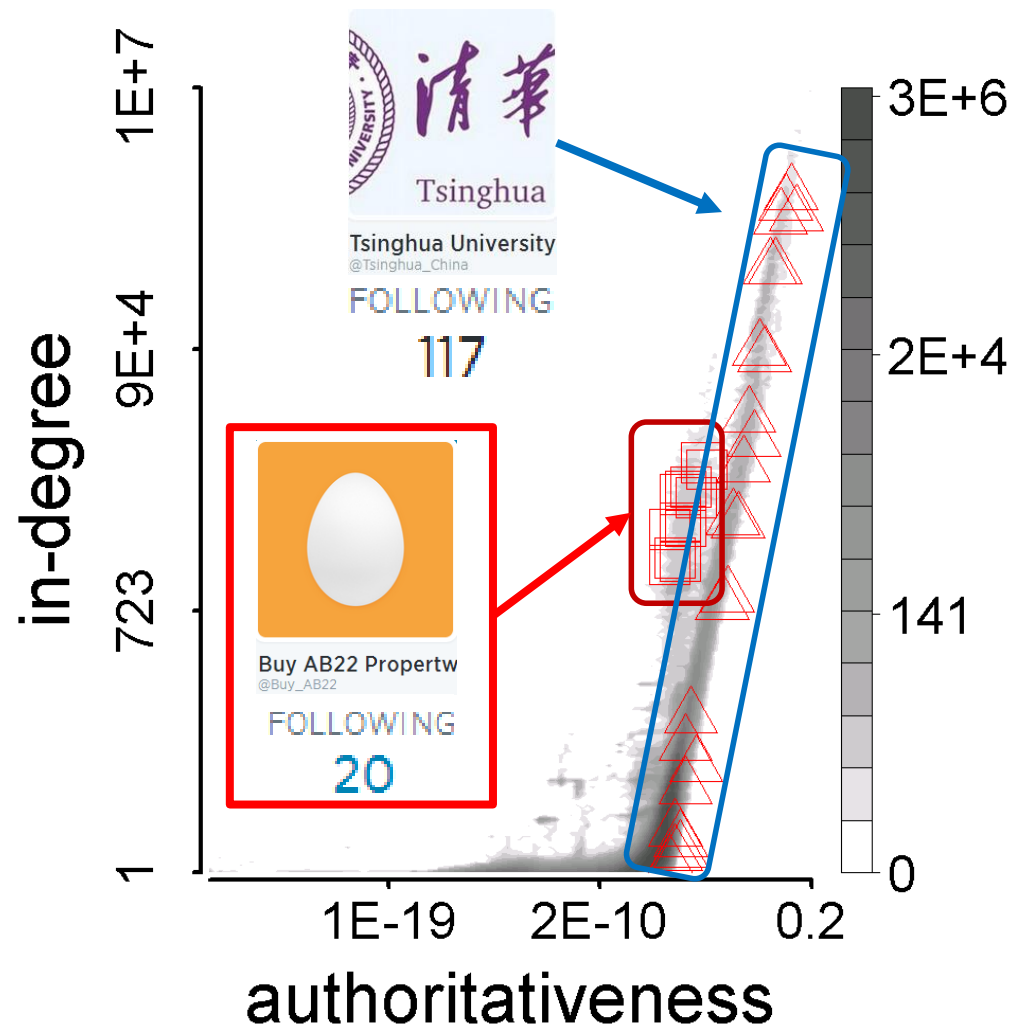
- Problem 5. **Synchronized behavior detection** (inferring zombie followers in Twitter networks) [PAKDD'14, KDD'14 best paper finalist, TKDD'15]



Synchronized Behavioral Patterns



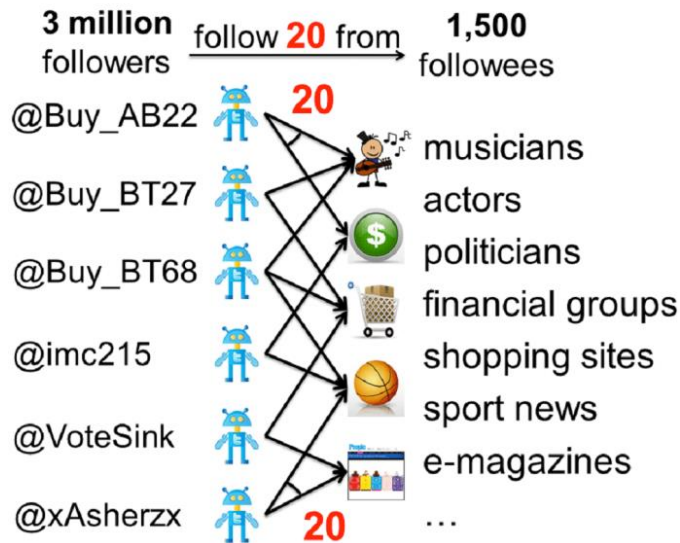
Synchronized Behavioral Patterns



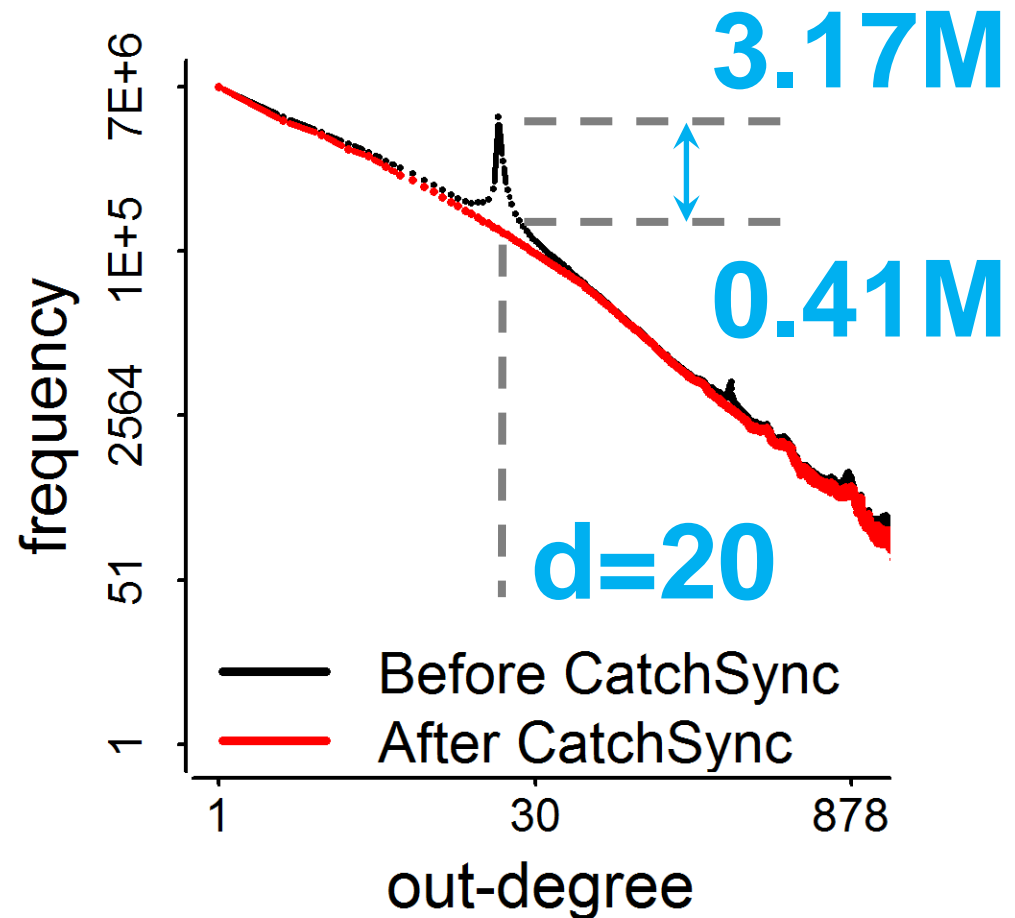
Recovering Distorted Out-degree Distribution



41M

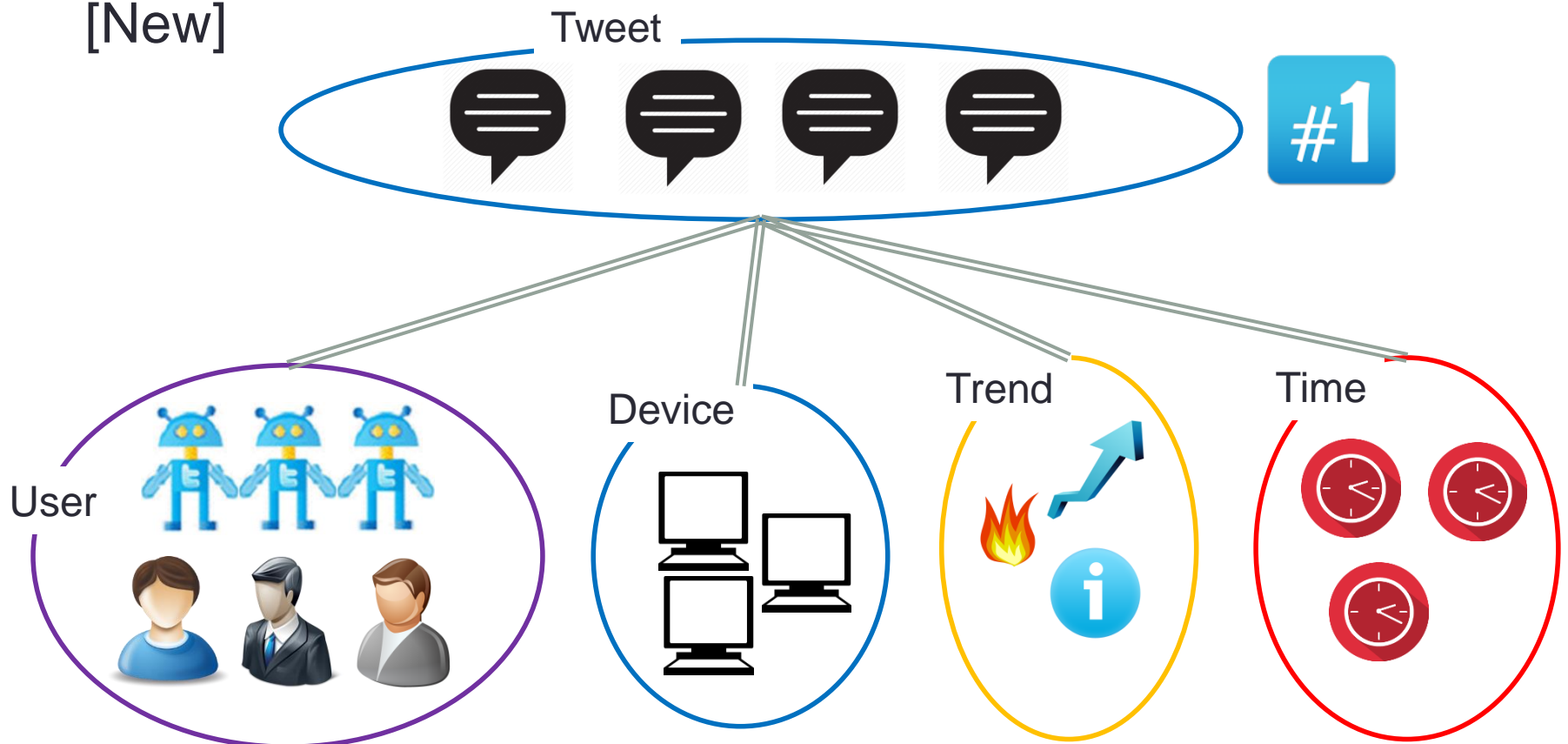


→ following behavior



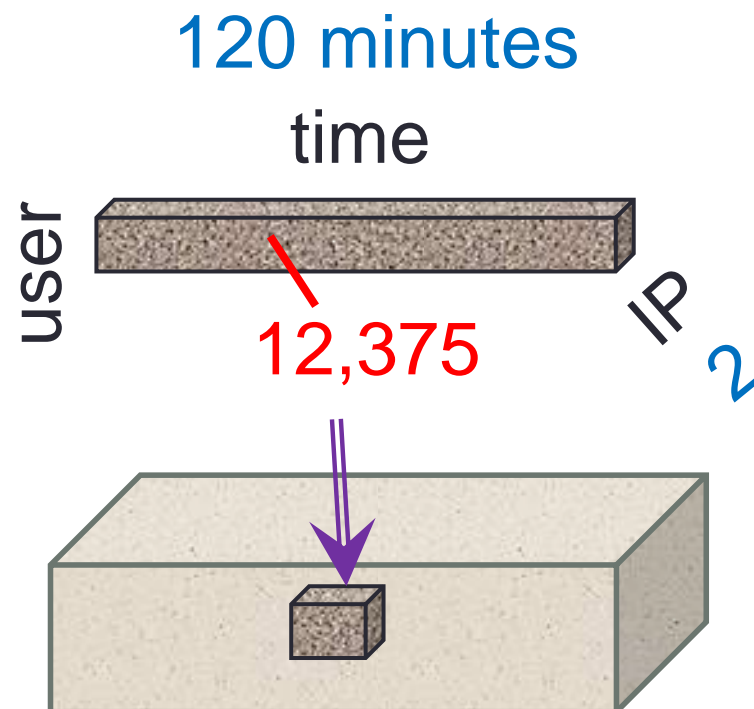
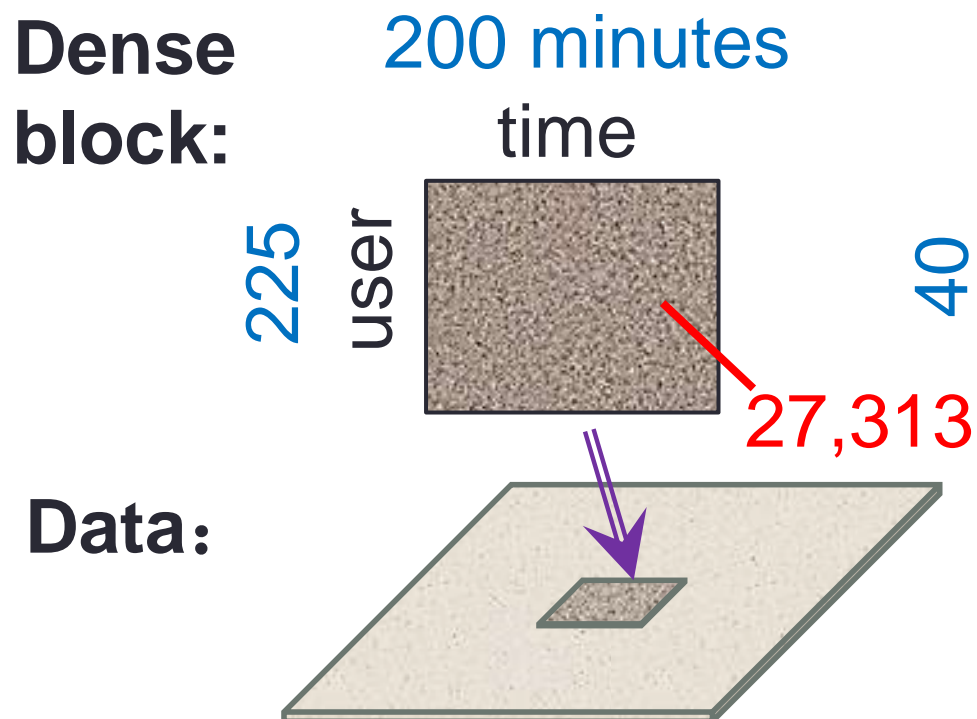
FROM Behavior Modeling TO Mining Behavioral Networks

- Problem 6. **Suspicious behavior detection** (catching astroturfing in Twitter retweeting and trending networks)
[New]



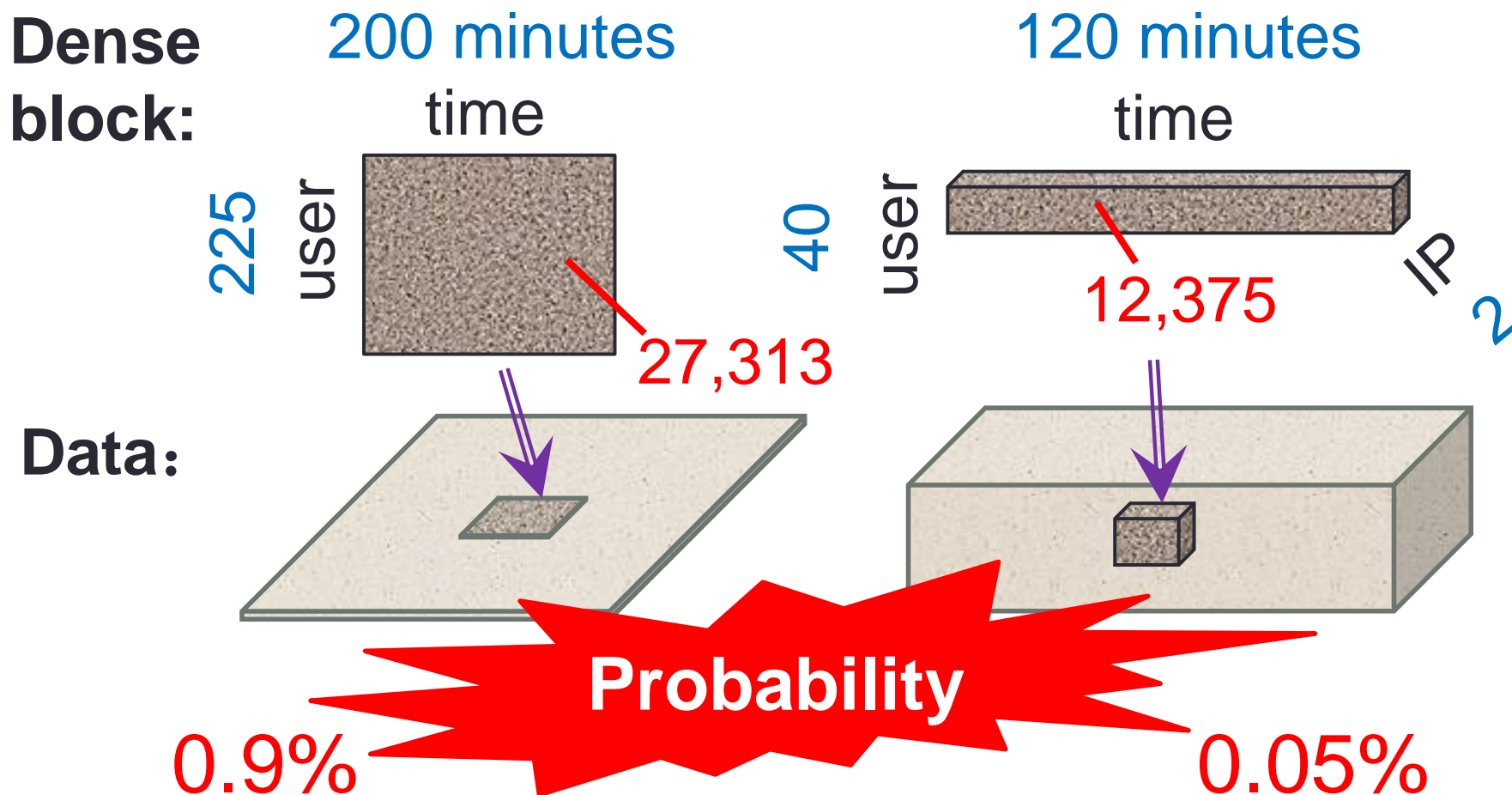
Measuring Suspiciousness: Multi-modal Dense Blocks

- 2 modes, and 3 modes, which is more suspicious?



Measuring Suspiciousness: Multi-modal Dense Blocks

- 2 modes, and 3 modes, which is more suspicious?



ROADMAP

Background

Previous works (6 problems)

Thinking & Future works

Challenges in Mining Behavioral Networks (Behavior Modeling)

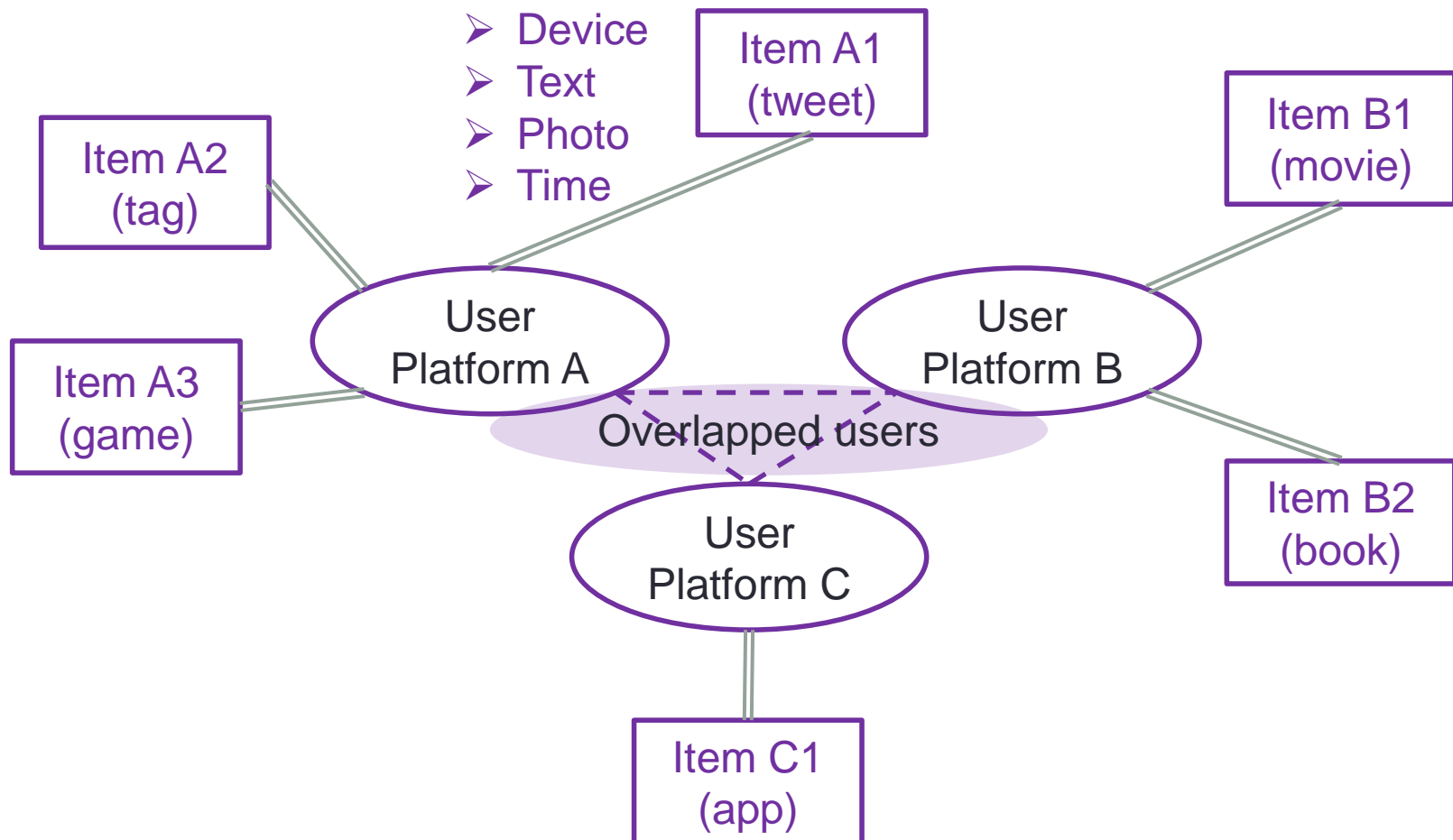
- **Sparse, massive, complex** behavioral *data*
 - Social, spatial, temporal contexts (Problem 1,2)
 - Walking across domains and platforms (Problem 3,4)
 - Anomalies/Suspiciousness (Problem 5,6)
- **Unified, structured, heterogeneous** *networks*
 - Construction and integration
 - OLAP and mining
 - Network-based anomaly detection
- **Effective, scalable** *algorithms*
- **Practical** *solutions*: We are changing users' behaviors.
 - Mining “plans”, “strategies” instead of “missing links”?

Construction and Integration of Behavioral Networks: Human-centered

■ Social networks

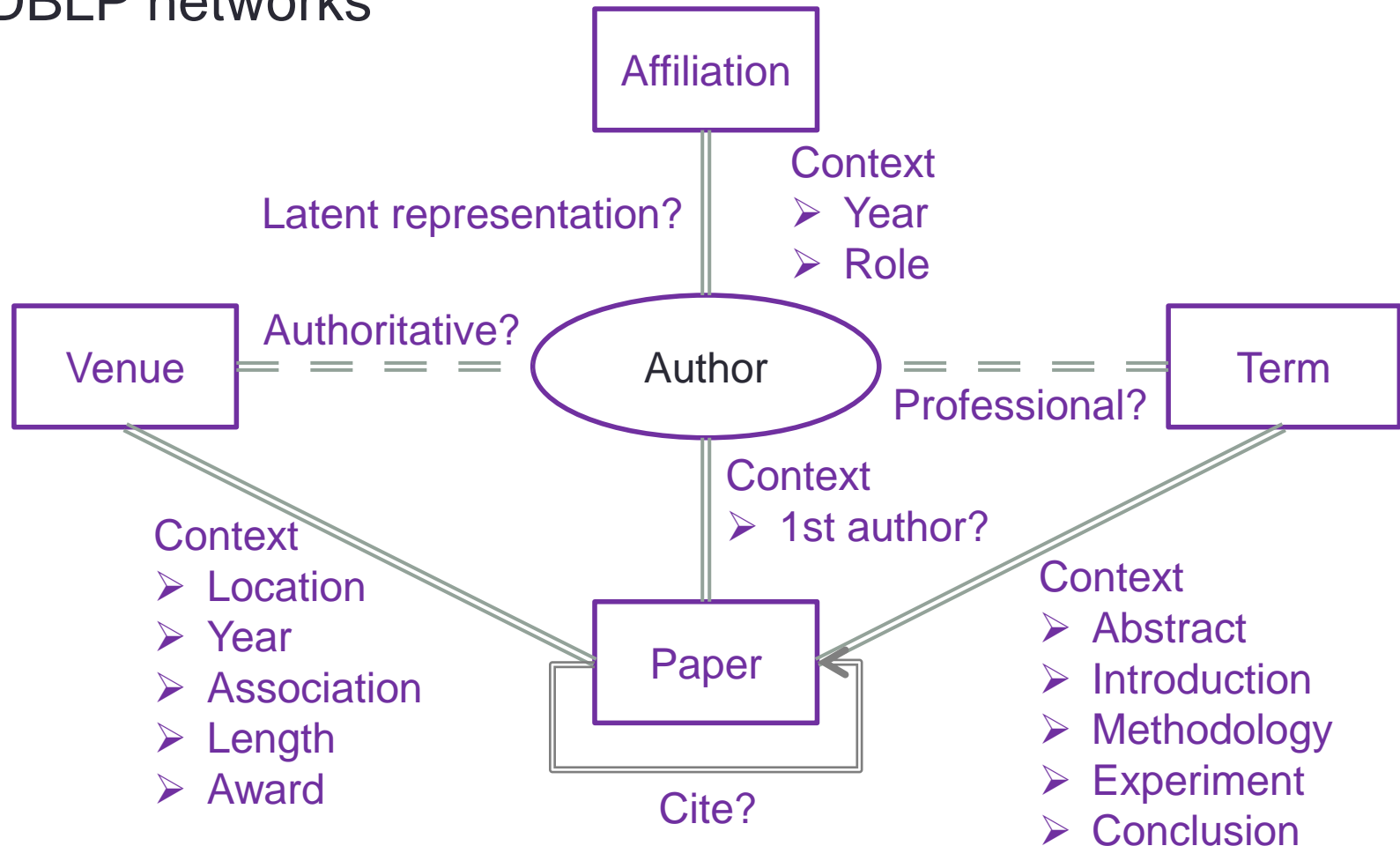
Context

- Location
- Device
- Text
- Photo
- Time



Construction and Integration of Behavioral Networks: Human-centered

■ DBLP networks



OLAP and Mining Behavioral Networks

■ Explainable citation recommendation (DBLP)

Affiliation

Author

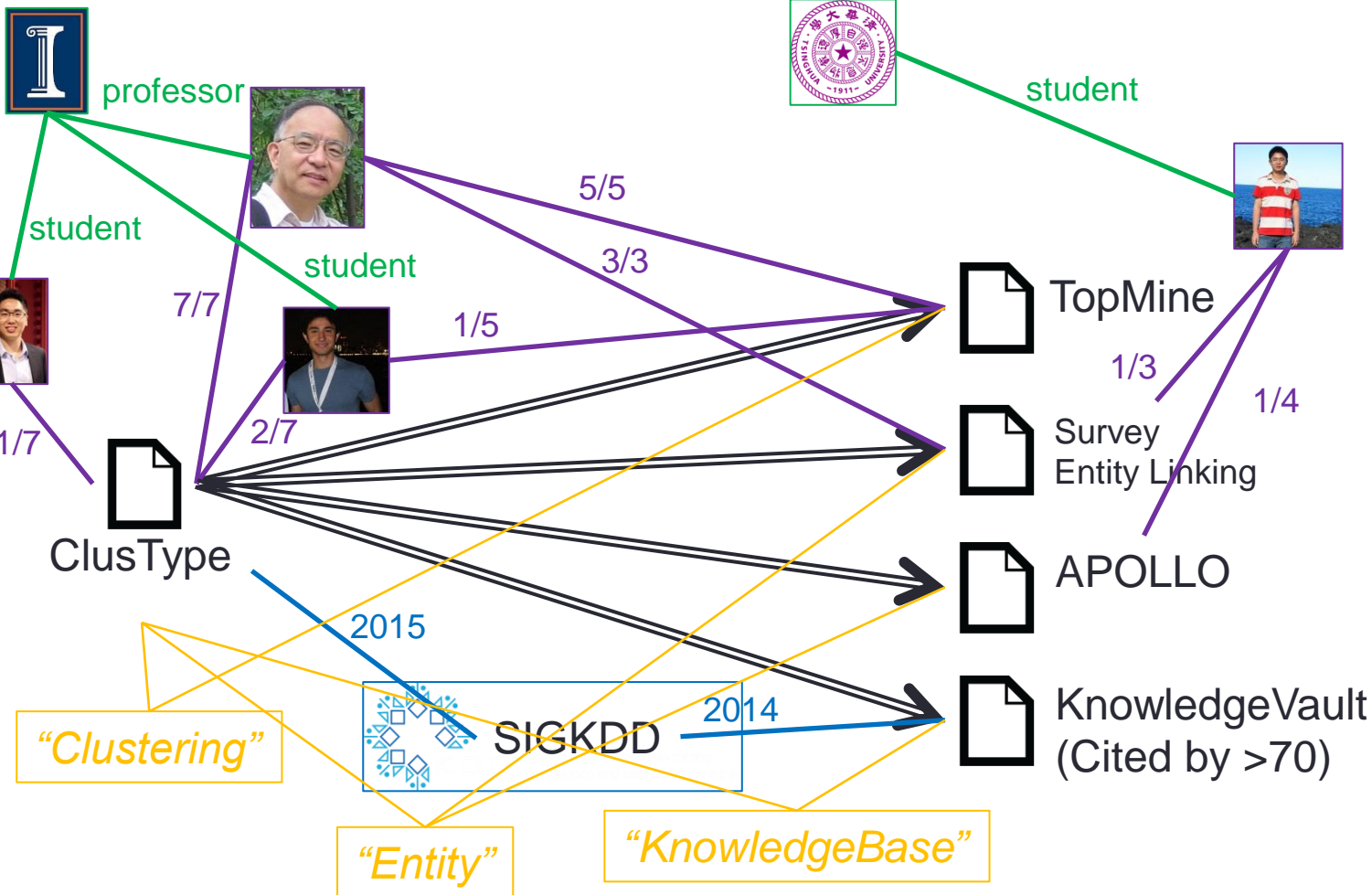
Paper

Venue

Term

One-step
explanation:

- Co-authored students'?
- Co-authored advisors'?
- Same affiliation?
- Same venues?
- Same terms?
- Well-cited paper? (Award?)



OLAP and Mining Behavioral Networks

■ Explainable submission recommendation (DBLP)

- For paper, recommending venues:
 - Paper-Author-Paper-Venue
 - Paper-Term-Paper-Venue
- For venue, recommending papers (“related to this conference?”):
 - Paper-Author-Paper-Venue
 - Paper-Term-Paper-Venue

■ Explainable content/item recommendation (Social media)

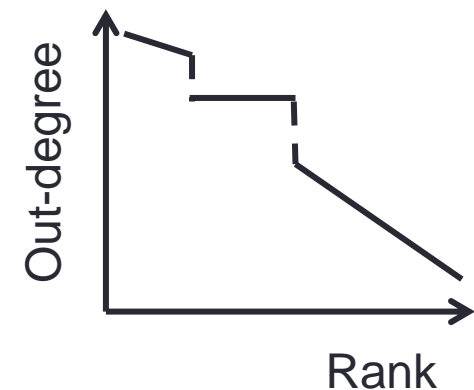
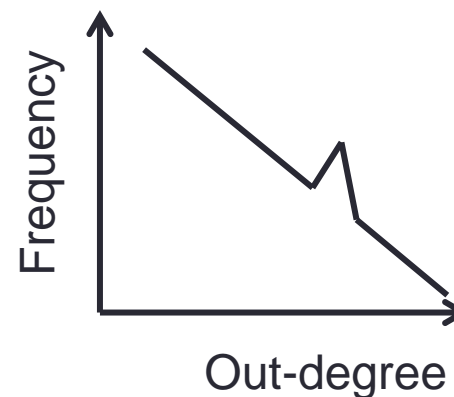
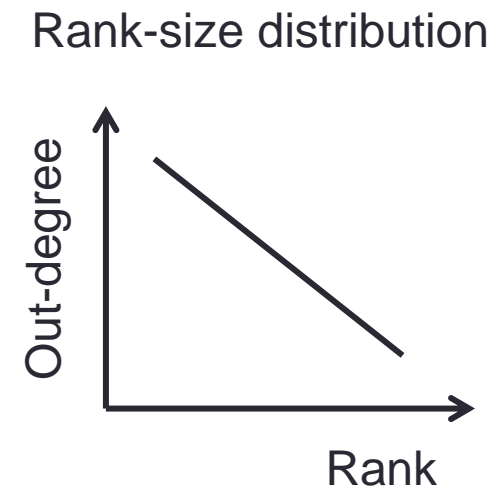
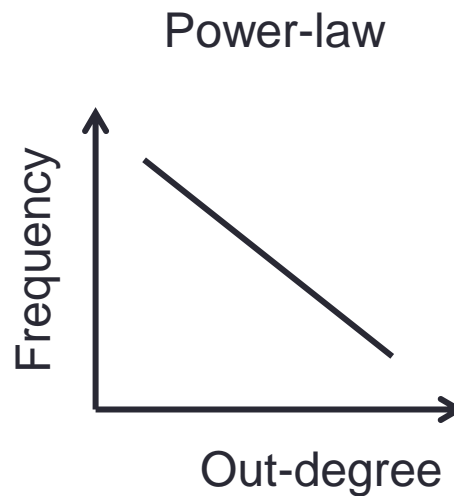
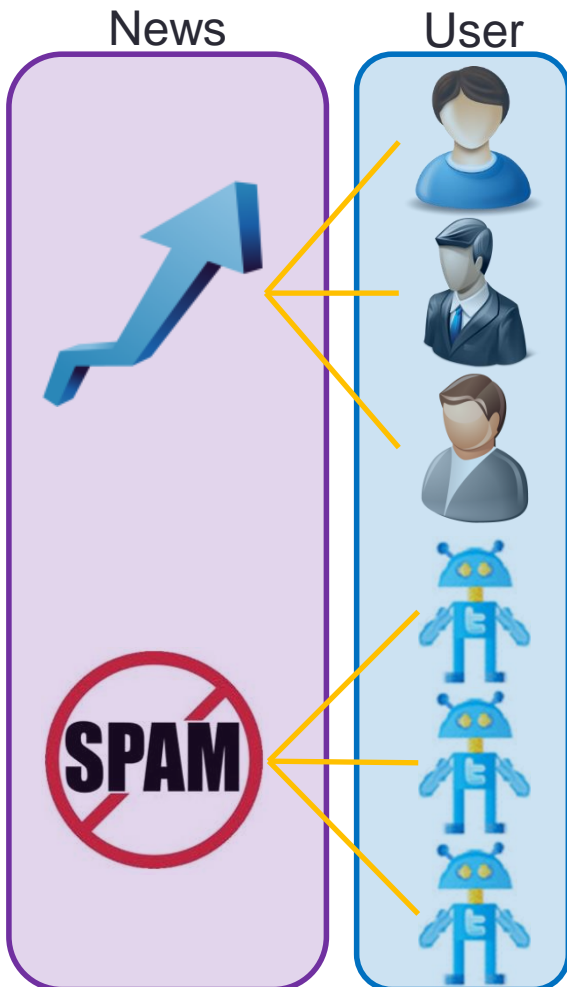
- User-ItemA-ItemA
- User-User-ItemA
- User-ItemB-ItemA
- User-ItemB-User-ItemA
- User-Context(Social, Spatial, Temporal, Domains, Platforms)-Item

■ Explainable APP recommendation (Tencent App Store)

- User-Context(?)-APP

Network-based Anomaly Detection

■ Natural outbreaks vs. Artificial promotions



Practical Solutions

■ Promotional strategy

- How to promote a policy/news/product (“ObamaCare”/“Rh-Blood”/“Earthquake”) on social media?
 - Given an item, a set of promoters, a network, contexts (time, etc.)
 - Find a set of strategies (what, how, when the promoters operate)
 - Maximize `#infected_users_who_adopt_the_item`
- How to promote a research topic/area/direction (“heterogeneous networks”, “cascade prediction”, “deep learning”, etc.) in academic communities?
 - Given a topic, a set of authors (students, friends, etc.), contexts (conferences, journals, time, etc.)
 - Propose a set of papers
 - Maximize `#infected_researchers_who_work_on_the_topic`

THANK YOU

References coming ...

References: Journal Papers

- **Meng Jiang**, Peng Cui, Fei Wang, Wenwu Zhu and Shiqiang Yang. “Social Recommendation with Cross-Domain Transferable Knowledge”, in IEEE TKDE 2015. (to appear. Regular. IF=1.815. CCF A)
- **Meng Jiang**, Peng Cui, Alex Beutel, Christos Faloutsos and Shiqiang Yang. “Catching Synchronized Behaviors in Large Networks: A Graph Mining Approach”, in ACM TKDD 2015. (to appear. Full. IF=1.147. CCF B)
- **Meng Jiang**, Peng Cui, Fei Wang, Wenwu Zhu and Shiqiang Yang. “Scalable Recommendation with Social Contextual Information”, in IEEE TKDE 2014. (Regular. IF=1.815. CCF A)
- Lu Liu, Feida Zhu, **Meng Jiang**, Jiawei Han, Lifeng Sun and Shiqiang Yang. “Mining Diversity on Social Media Networks”, in Multimedia Tools and Applications 2012. (CCF B)

References: Conference Papers

- **Meng Jiang**, Peng Cui, Alex Beutel, Christos Faloutsos and Shiqiang Yang. “CatchSync: Catching Synchronized Behavior in Large Directed Graph”, in ACM SIGKDD 2014. (Full. **Best paper finalist**. Acc. rate=14.6%. CCF A)
- **Meng Jiang**, Peng Cui, Fei Wang, Xinran Xu, Wenwu Zhu and Shiqiang Yang. “FEMA: Flexible Evolutionary Multi-faceted Analysis for Dynamic Behavioral Pattern Discovery”, in ACM SIGKDD 2014. (Full. Acc. rate=14.6%. CCF A)
- **Meng Jiang**, Peng Cui, Alex Beutel, Christos Faloutsos and Shiqiang Yang. “Inferring Strange Behavior from Connectivity Pattern in Social Networks”, in PAKDD 2014. (Full. Acc. rate=10.8%. CCF C)
- **Meng Jiang**, Peng Cui, Alex Beutel, Christos Faloutsos and Shiqiang Yang. “Detecting Suspicious Following Behavior in Multimillion-Node Social Networks”, in WWW 2014. (*Poster*. CCF B)

References: Conference Papers

- **Meng Jiang**, Peng Cui, Rui Liu, Qiang Yang, Fei Wang, Wenwu Zhu and Shiqiang Yang. “Social Contextual Recommendation”, in CIKM 2012. (Full. Acc. rate=13.4%. 65 citations till 06/2015. CCF B)
- **Meng Jiang**, Peng Cui, Fei Wang, Qiang Yang, Wenwu Zhu and Shiqiang Yang. “Social Recommendation across Multiple Relational Domains”, in CIKM 2012. (Full. Acc. rate=13.4%. 26 citations till 06/2015. CCF B)
- Lu Liu, Jie Tang, Jiawei Han, **Meng Jiang** and Shiqiang Yang. “Mining Topic-Level Influence in Heterogeneous Networks”, in CIKM 2010. (CCF B)

References: Submitted Papers

- **Meng Jiang**, Peng Cui, and Christos Faloutsos. “Suspicious Behavior Detection: Current Trends and Future Directions”, to IEEE Intelligent Systems Magazine Special Issue on Online Behavioral Analysis and Modeling.
- **Meng Jiang**, Alex Beutel, Peng Cui, Bryan Hooi, Shiqiang Yang, and Christos Faloutsos. “A General Suspiciousness Metric for Dense Blocks in Multi-Modal Data”, to IEEE ICDM, 2015.
- **Meng Jiang**, Peng Cui, Nicholas Jing Yuan, Xing Xie, and Shiqiang Yang. “Little is Much: Bridging Cross-Platform Behaviors Through Small Overlapped Crowds”, to IEEE ICDM, 2015.
- **Meng Jiang**, Peng Cui, Alex Beutel, Christos Faloutsos and Shiqiang Yang. “Inferring Lockstep Behavior from Connectivity Pattern in Large Graphs”, to Knowledge and Information Systems (KAIS).