Computational Models for Social Influence and Diffusion

Yang Yang and Jie Tang
Zhejiang University
Tsinghua University
Homepage: http://yangy.org   http://keg.cs.tsinghua.edu.cn/jietang/

Part I:
Learning User Behavior Influence in Large-Scale Social Networks

Networked World

facebook

- **1.65 billion** MAU
- **2.5 trillion** minutes/month

twitter

- **255 million** MAU
- **Peak: 143K** tweets/s

amazon.com

- **304 million** active users
- **14 billion** items/year

WeChat

- **QQ: 800 million** MAU
- **WeChat: 700 million** MAU

新浪微博 weibo.com

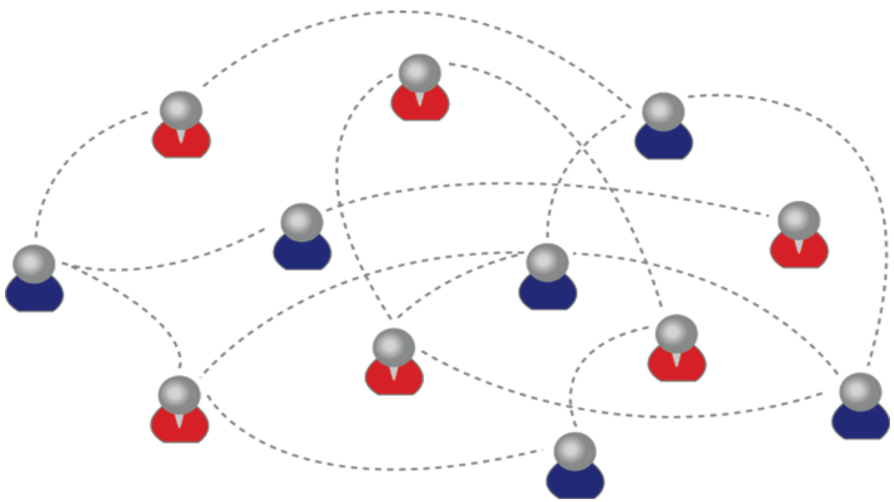
- **220 million** users
- **influencing** our daily life

Alibaba Group 阿里巴巴集团

- **~700 million** trans. (alipay)
- **120.7 billion** on 11/11

What is a social network?

- A social network is:
  - a **graph** made up of :
  - a set of **individuals**, called “nodes”, and
  - tied by one or more **interdependency**, such as friendship, called “edges”.



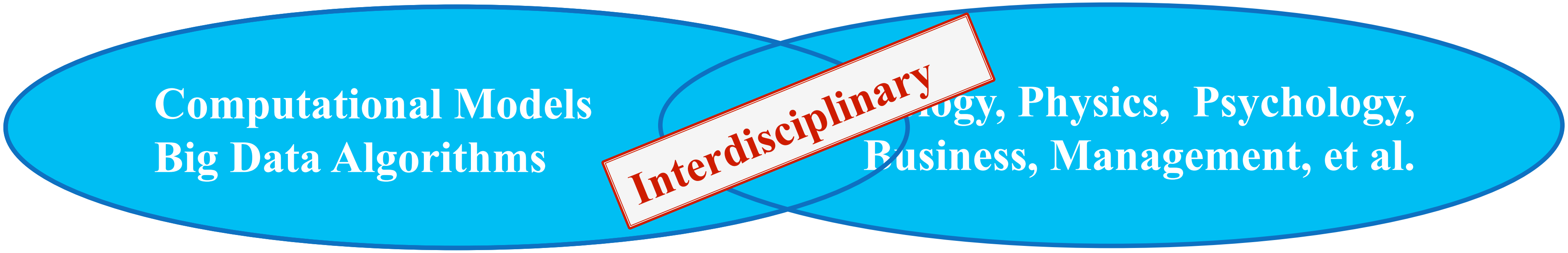
Computational Social Science

Computational Social Science [Giles] 2012

Computational Social Science [Lazer et al.] 2009

"A field is emerging that leverages the capacity to collect and analyze **data at a scale** that may reveal patterns of *individual* and *group behaviors*."

David Lazer, Alex Pentland, Lada Adamic, Sinan Aral, Alber-Laszlo Barabasi, et al. from *Departments of Sociology, Computer Science, Physics, Business, Government*, etc. at *Harvard, MIT, Northeastern, Northwestern, Columbia, Cornell, etc.*



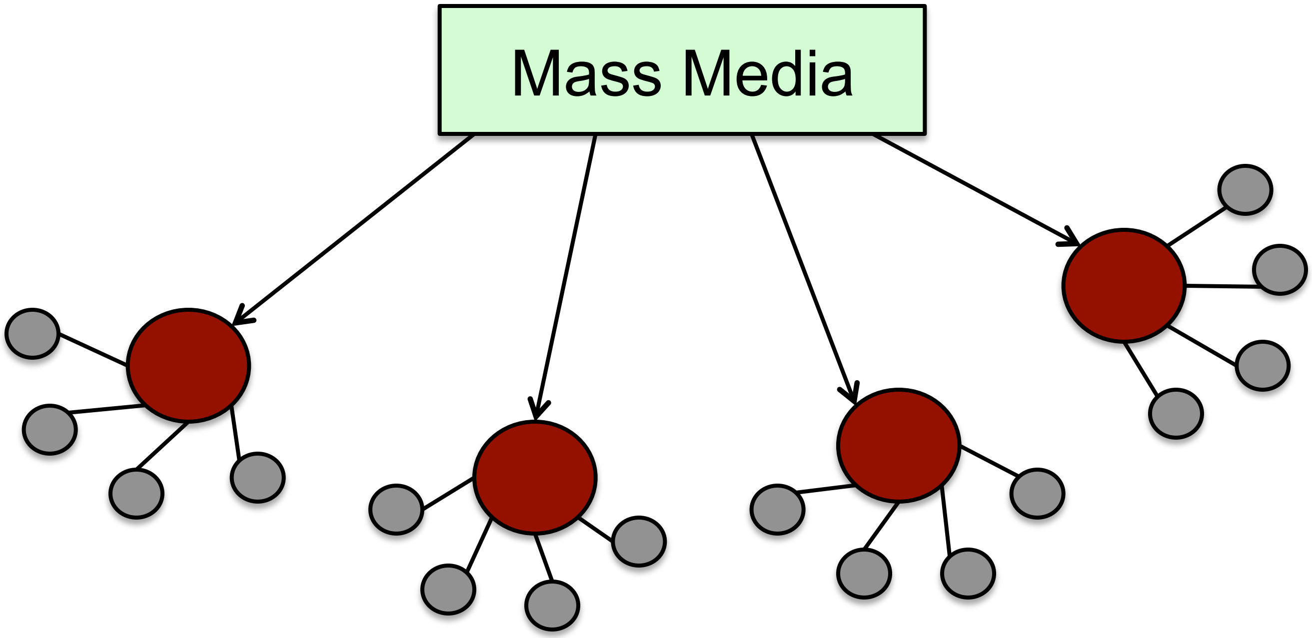
1. **David Lazer et al. Computational Social Science. *Science* 2009.**
2. James Giles. Computational Social Science: Making the Links. **Nature** 2012.

What is Social Influence?

- Social influence occurs when one's opinions, emotions, or behaviors are affected by others, intentionally or unintentionally.[1]
  - Peer Pressure
  - Opinion leadership
  - Conformity
  - …

[1] http://en.wikipedia.org/wiki/Social_influence

Two-step Flow Theory

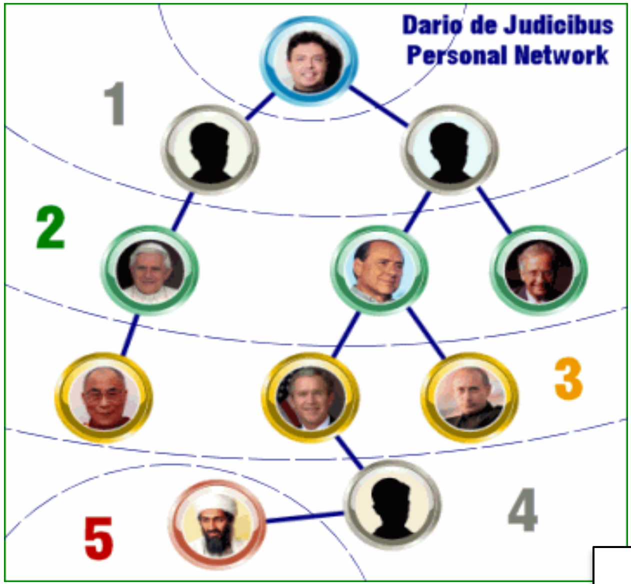


Opinion leader

Individuals in social contact with an opinion leader

The theory of "Three Degree of Influence"

Six degree of separation[1]



Three degree of Influence[2]

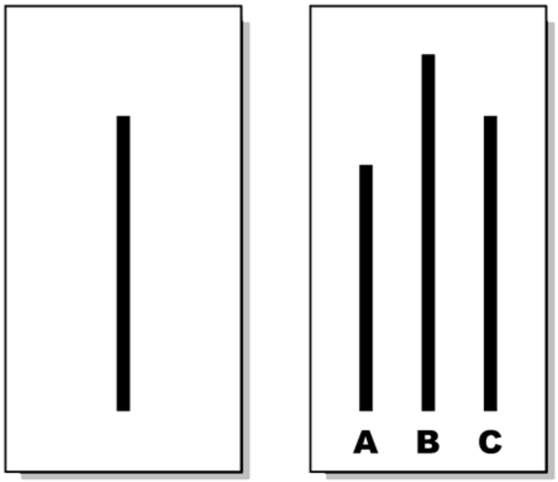
You are able to **influence** up to >1,000,000 persons in the world, according to the Dunbar's number[3].

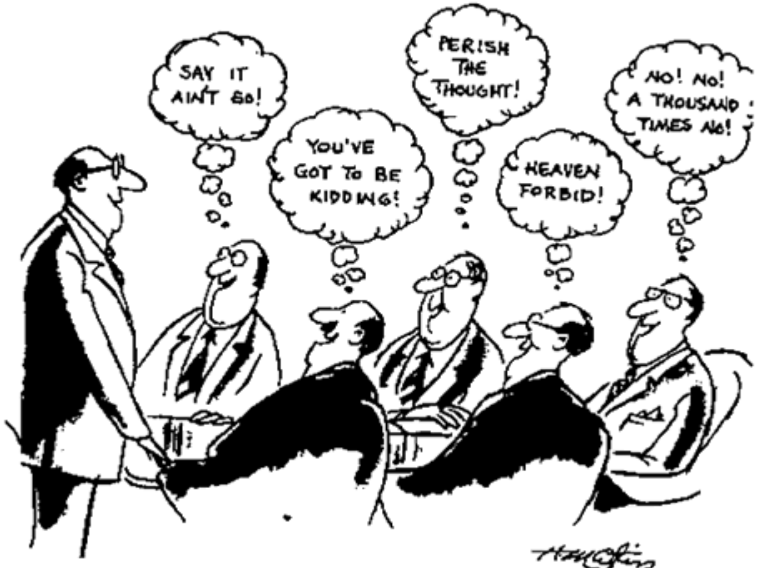
[1] S. Milgram. The Small World Problem. Psychology Today, 1967, Vol. 2, 60–67

[2] J.H. Fowler and N.A. Christakis. The Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis Over 20 Years in the Framingham Heart Study. British Medical Journal 2008; 337: a2338

[3] R. Dunbar. Neocortex size as a constraint on group size in primates. Human Evolution, 1992, 20: 469–493.

Asch's Experiment





"All those in favour say 'Aye'."

"Aye." "Aye." "Aye." "Aye." "Aye."

Which line matches the first line, A, B, or C?

74% of the participants followed the majority judgment on at least one trial, even when the majority was wrong.

Does Social Influence Really Matter?

- Case 1: Social influence and political mobilization[1]
  - Will online political mobilization really work?

A controlled trial (with 61M users on FB)

- Social msg group: was shown with msg that indicates one's friends who have made the votes.
- Informational msg group: was shown with msg that indicates how many other.
- Control group: did not receive any msg.

[1] R. M. Bond, C. J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle and J. H. Fowler. A 61-million-person experiment in social influence and political mobilization. Nature, 489:295-298, 2012.

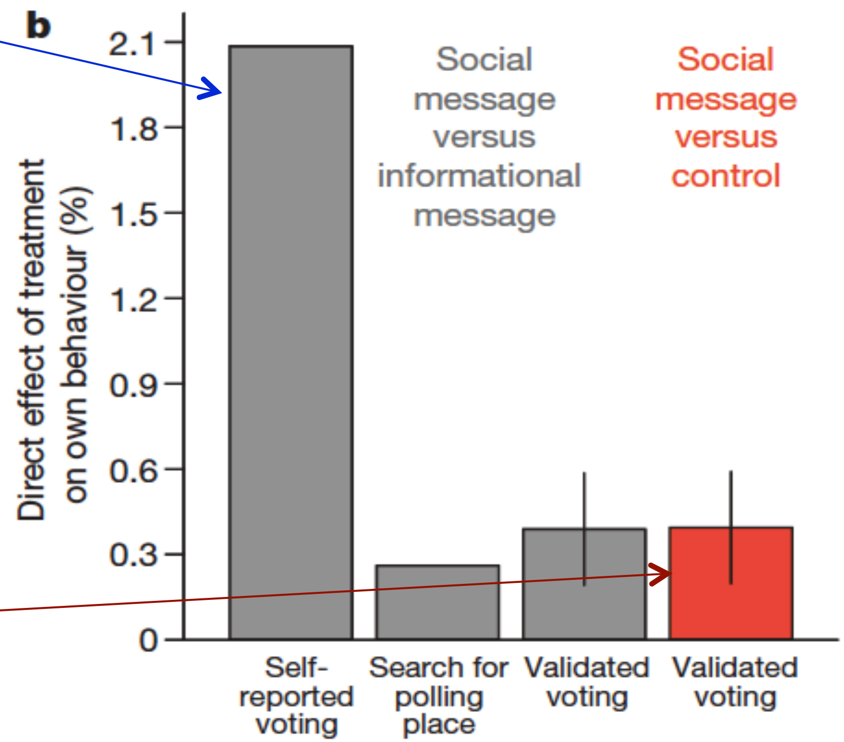
Does Social Influence Really Matter?

Social msg group v.s.
Info msg group

Result: The former were 2.08% (*t*-test, *P*<0.01) more likely to click on the "I Voted" button

Social msg group v.s.
Control group

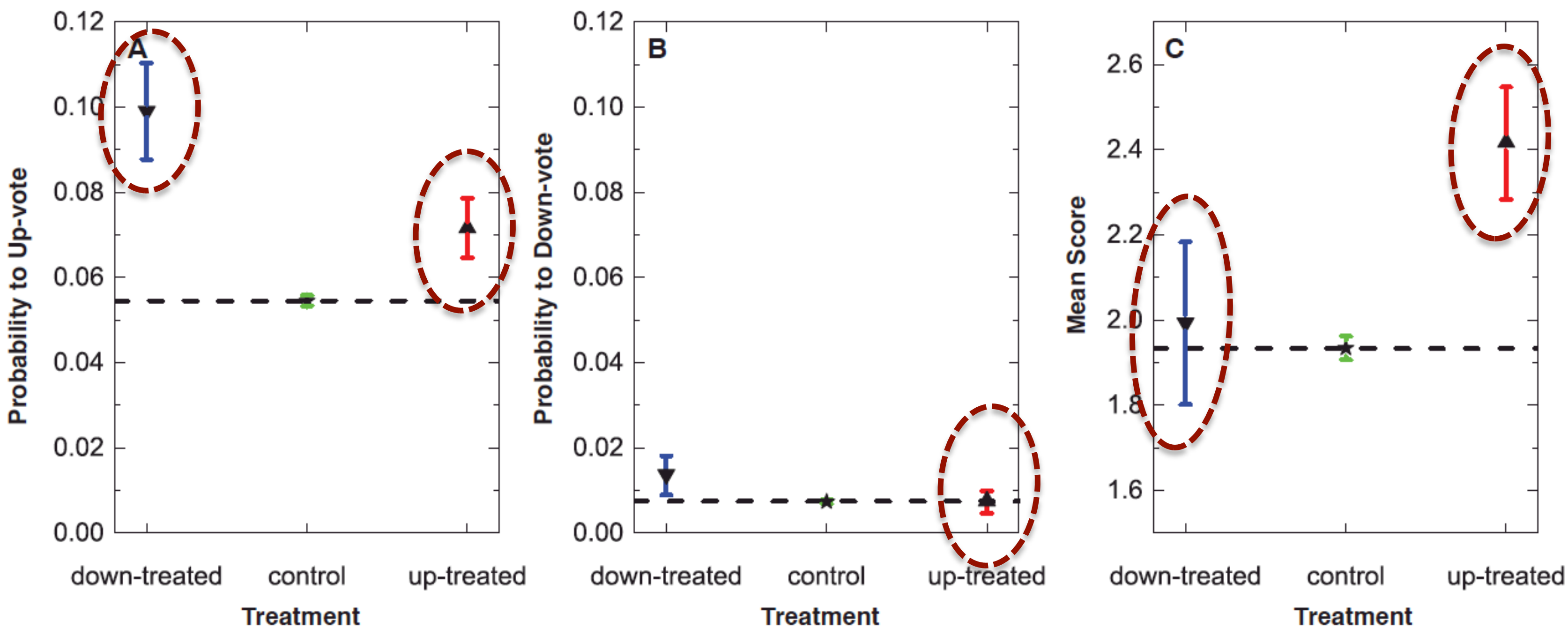
Result: The former were 0.39% (*t*-test, *P*=0.02) more likely to **actually vote** (via examination of public voting records)



[1] R. M. Bond, C. J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle and J. H. Fowler. A 61-million-person experiment in social influence and political mobilization. Nature, 489:295-298, 2012.

Does Social Influence Really Matter?

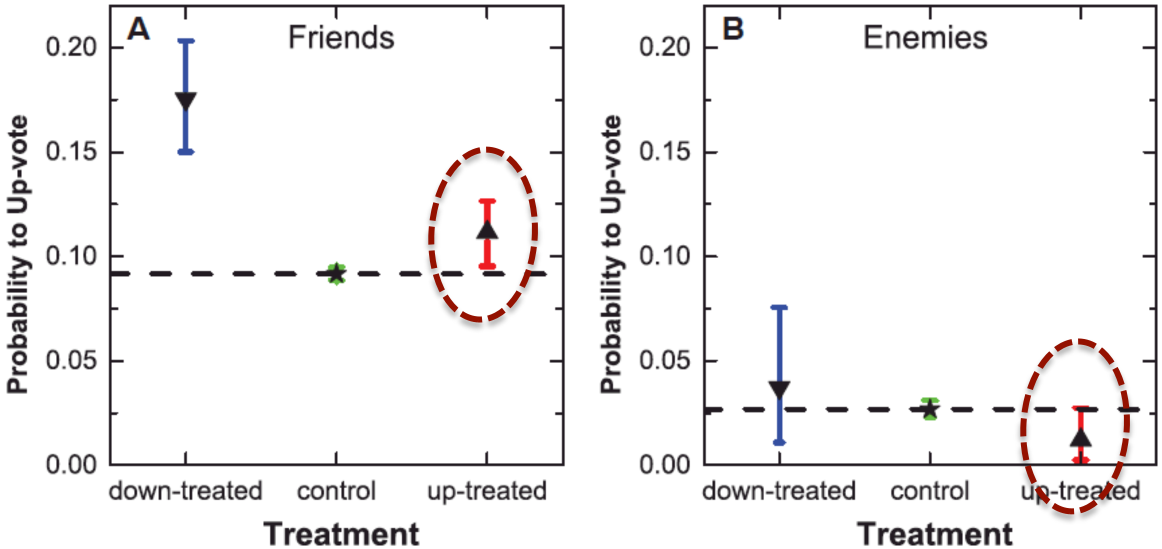
- Case 2: Social influence distorts decision-making [1]
  - Two treatment groups and a control group:
    - Up-treated: comments were artificially given a +1 rating;
    - Down-treated: comments were given a -1 rating;



[1] L. Muchnik, S. Aral, S. J. Taylor. Social Influence Bias: A Randomized Experiment. Science, Vol. 341, Issue 6146, page 647-651, 2013.

Does Social Influence Really Matter?

- Case 2: Social influence distorts decision-making [1]
  - Define a user's "friends" and "enemies" according to they "like" or "dislike" her (a feature of the studied web site)
  - Friendship moderates the impact of social influence.



Friends were more likely to up-vote a comment than enemies (9.2% versus 2.7%).

Friends tend to herd on current positive ratings (0.122 versus 0.092).

[1] L. Muchnik, S. Aral, S. J. Taylor. Social Influence Bias: A Randomized Experiment. Science, Vol. 341, Issue 6146, page 647-651, 2013.

We applied social influence to help real applications

—in very big Tencent networks

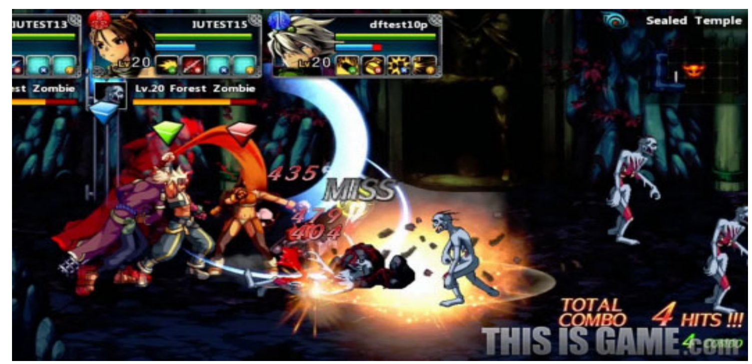
Big Data Analytics in Game Data

- Online gaming is one of the largest industries on the Internet…
- Facebook
  - 250 million users play games monthly
  - 200 games with more than 1 million active users
  - 12% of the company's revenue is from games
- Tencent (Market Cap: ~150B $)
  - More than 400 million gaming users
  - 50% of Tencent's overall revenue is from games

[1] Zhanpeng Fang, Xinyu Zhou, Jie Tang, Wei Shao, A.C.M. Fong, Longjun Sun, Ying Ding, Ling Zhou, and Jarder Luo. Modeling Paying Behavior in Online Social Networks. **CIKM'14.**

Two games: DNF

- Dungeon & Fighter Online (DNF)
  - A game of melee combat between users and large number of underpowered enemies
  - 400+ million users, the 2^{nd} largest online game in China
  - Users in the game can fight against enemies by individuals or by groups



Two games: QQ Speed

- QQ Speed
  - A racing game that users can partake in competitions to play against other users
  - 200+ million users
  - Users can race against other users by individuals or forma a group to race together
  - Some users may pay…

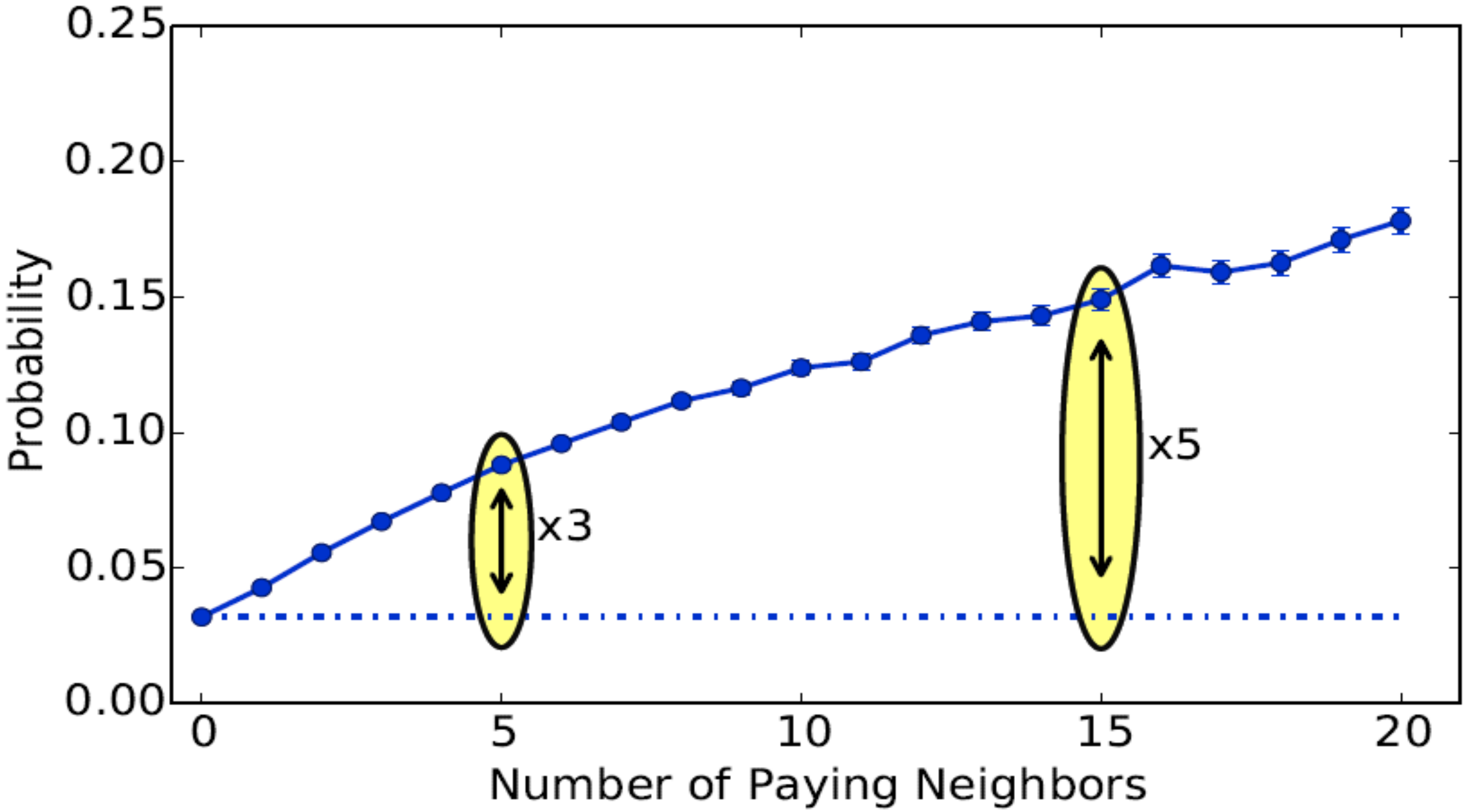
Task

- Given behavior log data and paying logs of online game users, predict

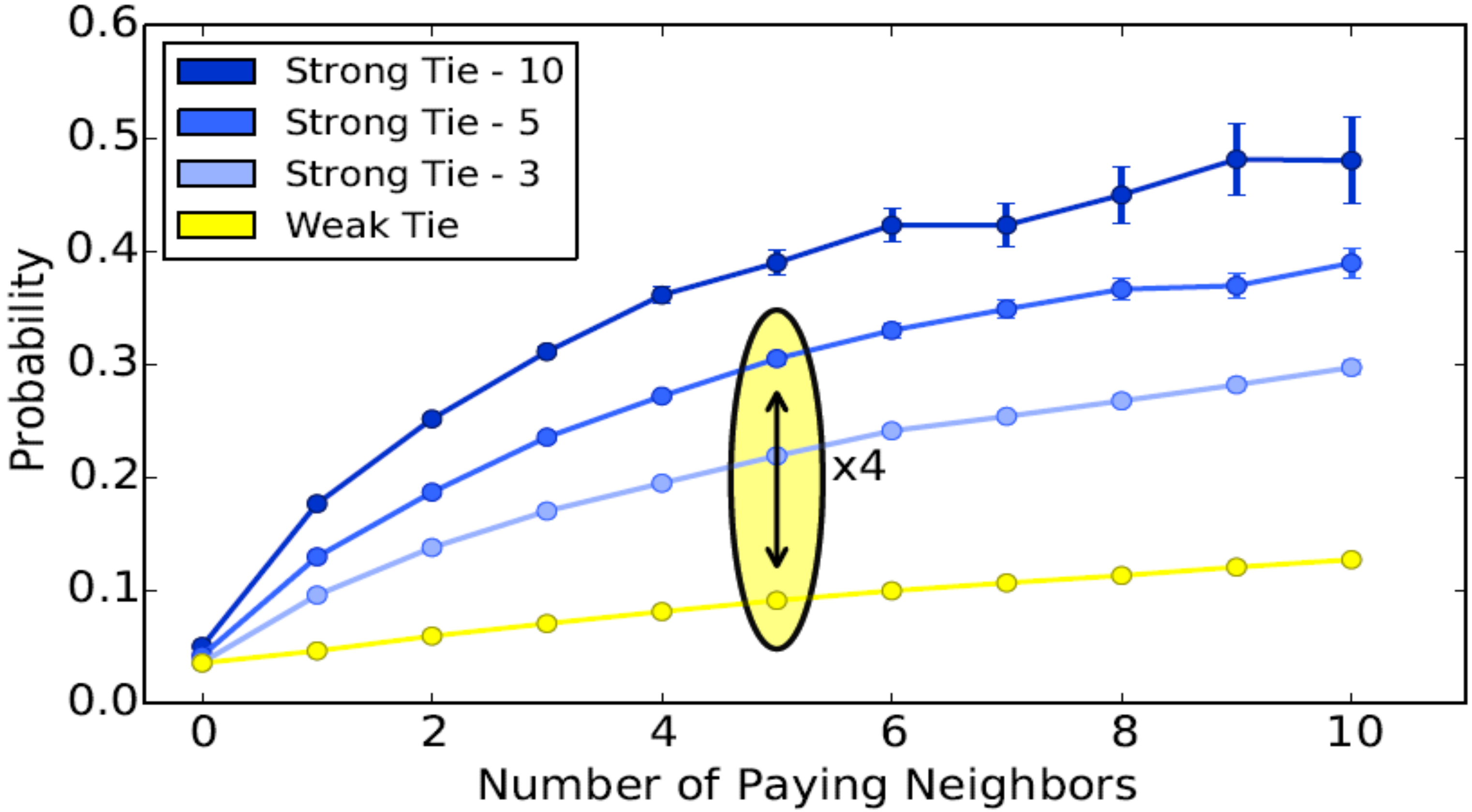
Free users -> Paying users

- Will social influence play an important role in this task?

Social Influence

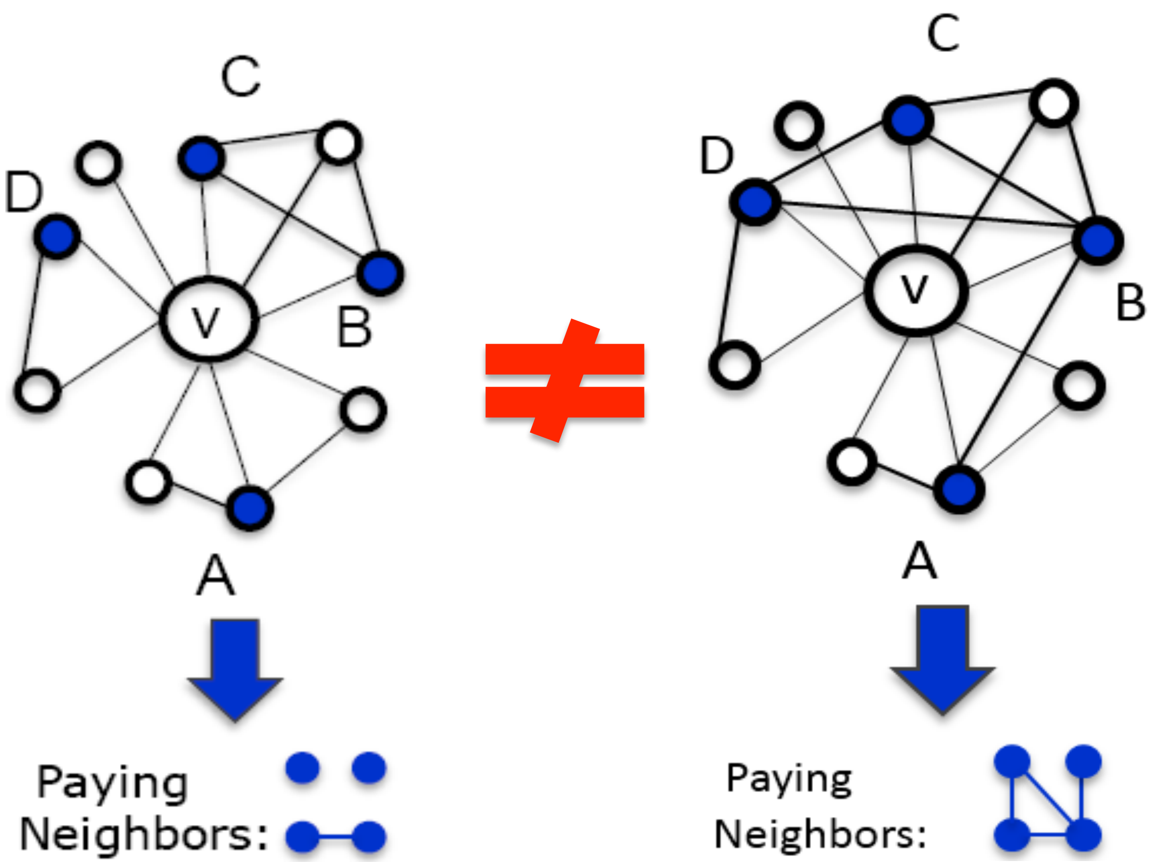


Influence + Tie Strength



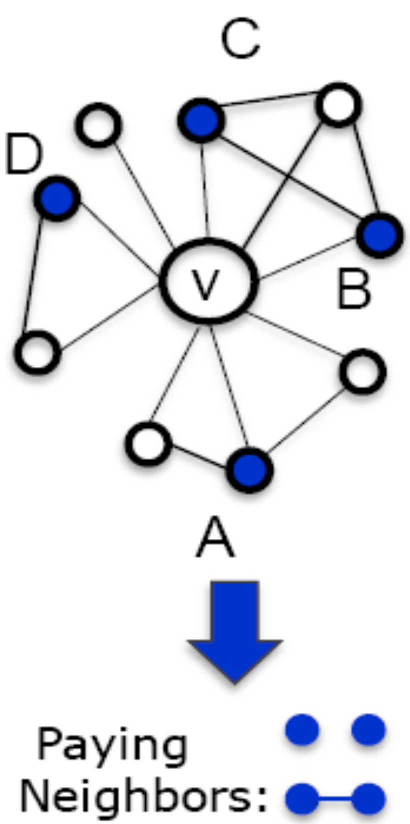
Structure Diversity

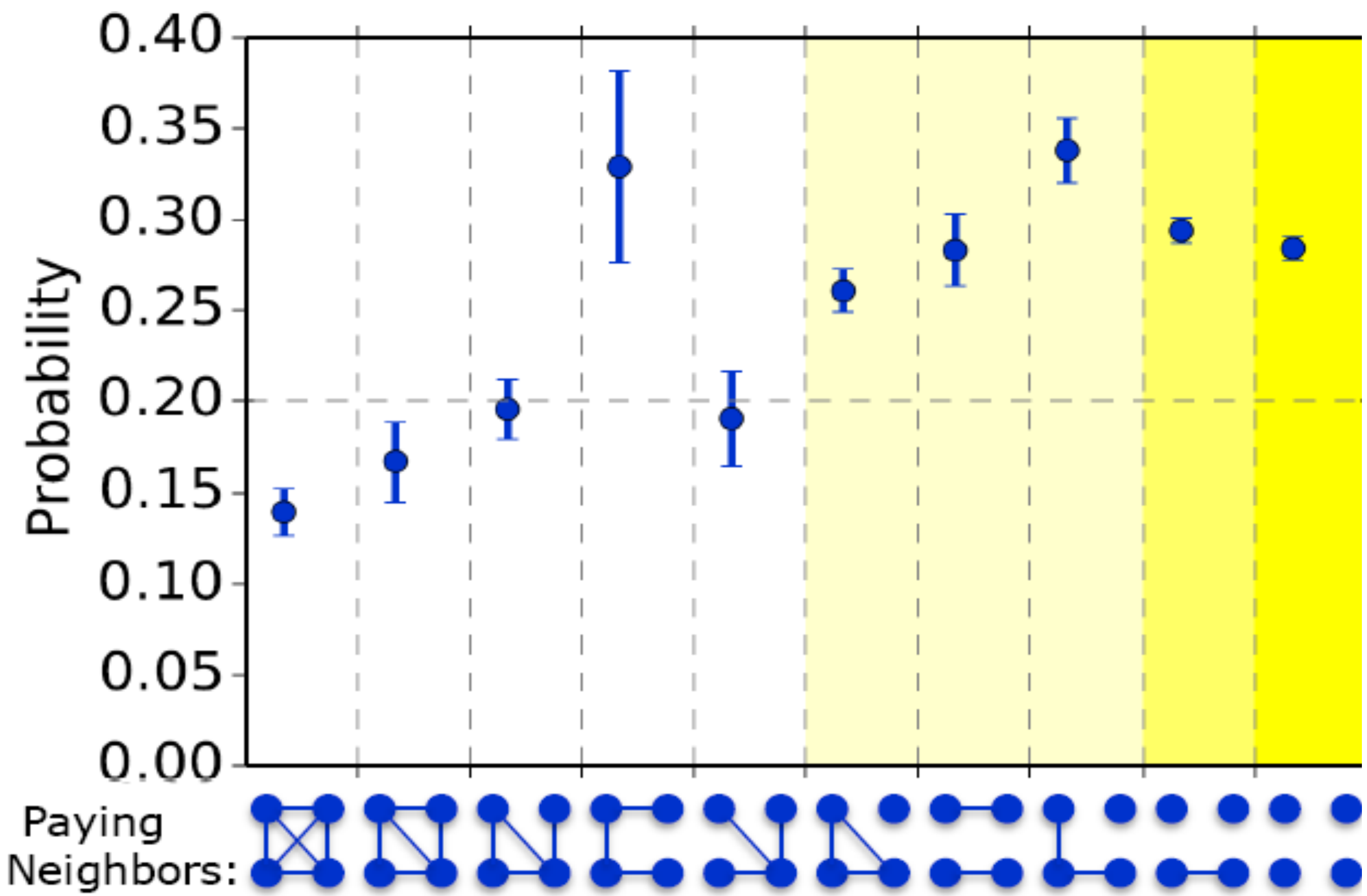
Different structures of a user's neighbors have different effects on the user's behavior[1]



[1] Ugander, J., Backstrom, L., Marlow, C., & Kleinberg, J. Structural diversity in social contagion. In PNAS'12.

Structure Diversity





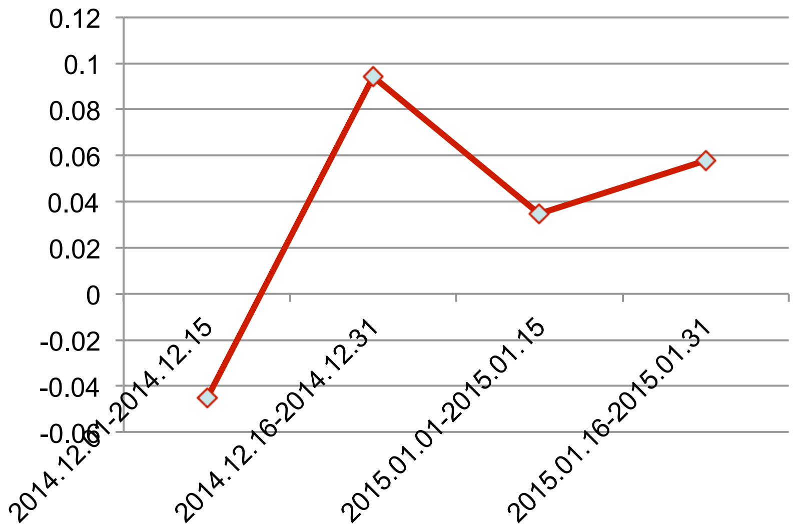
Online Test

- Test setting
  - Two groups: *test group* and *control group*
  - Send msgs to invite the user to attend a promotion activity.

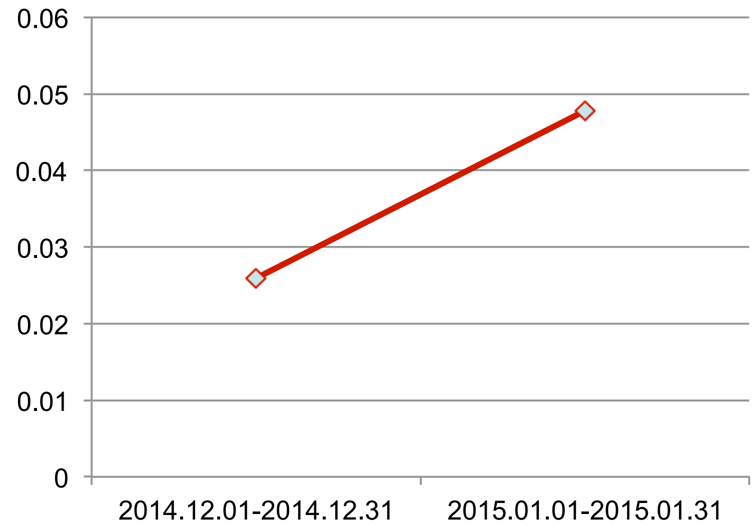
	Online Test 1 2013.12.27 - 2014.1.3		Online Test 2 2014.1.24 - 2014.1.27		
Group name	test group	control group	test group	test group2	control group
Group size	600K	200K	400K	400K	200K
#Message read Message read rate	345K 57.50%	106K 53.00%	229K 57.25%	215K 53.75%	106K 53.00%
#Message clicked Message clicked rate	47584 7.93%	7466 3.73%	23325 5.83%	20922 5.23%	6299 3.15%
Lift_Ratio	196.87%	0%	123.63%	73.40%	0%

Online Test

- Item Recommendation



Half-Month Improvement



Single-Month Improvement

Our social influence based recommendation algorithm in QQ Speed increased the item income by **9.4%** during December, 2014.

How to Model the Diffusion of Social Influence in Networks?

Compartmental Models in Epidemiology

- The **SIR** model, which is proposed by Kermack and McKendrick in the early 1900s.
- The model predicts infectious diseases

- Transition rates:

$$\frac{dS}{dt} = -\beta S(t)I(t)$$
$$\frac{dI}{dt} = \beta S(t)I(t) - \gamma I(t)$$
$$\frac{dR}{dt} = \gamma I(t)$$

S(t) : susceptible individuals at time t;

I(t) : infected individuals at time t;

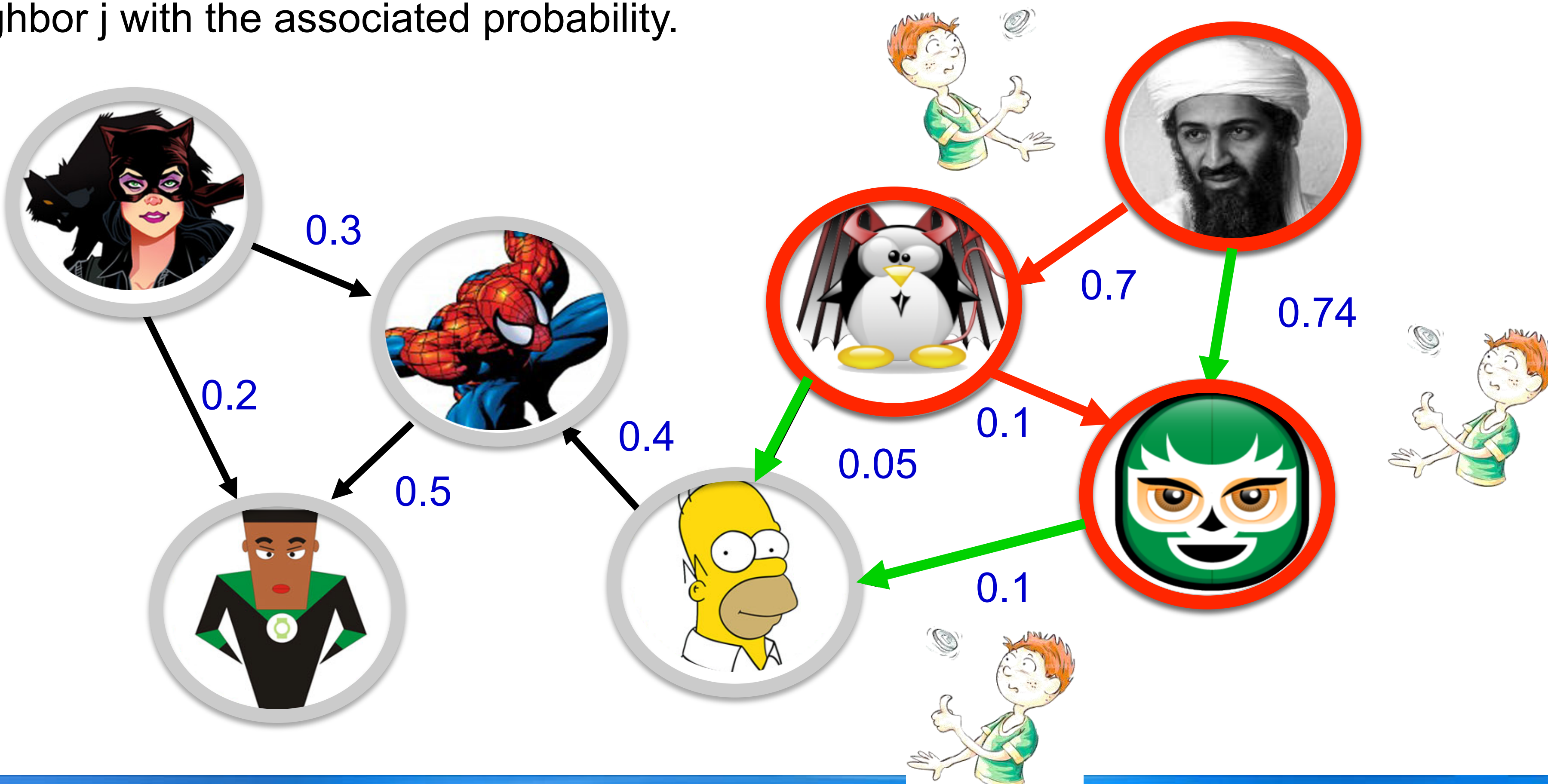
R(t) : recovered individuals at t ;

β *:* the contact rate;

γ *:* rate of recovery.

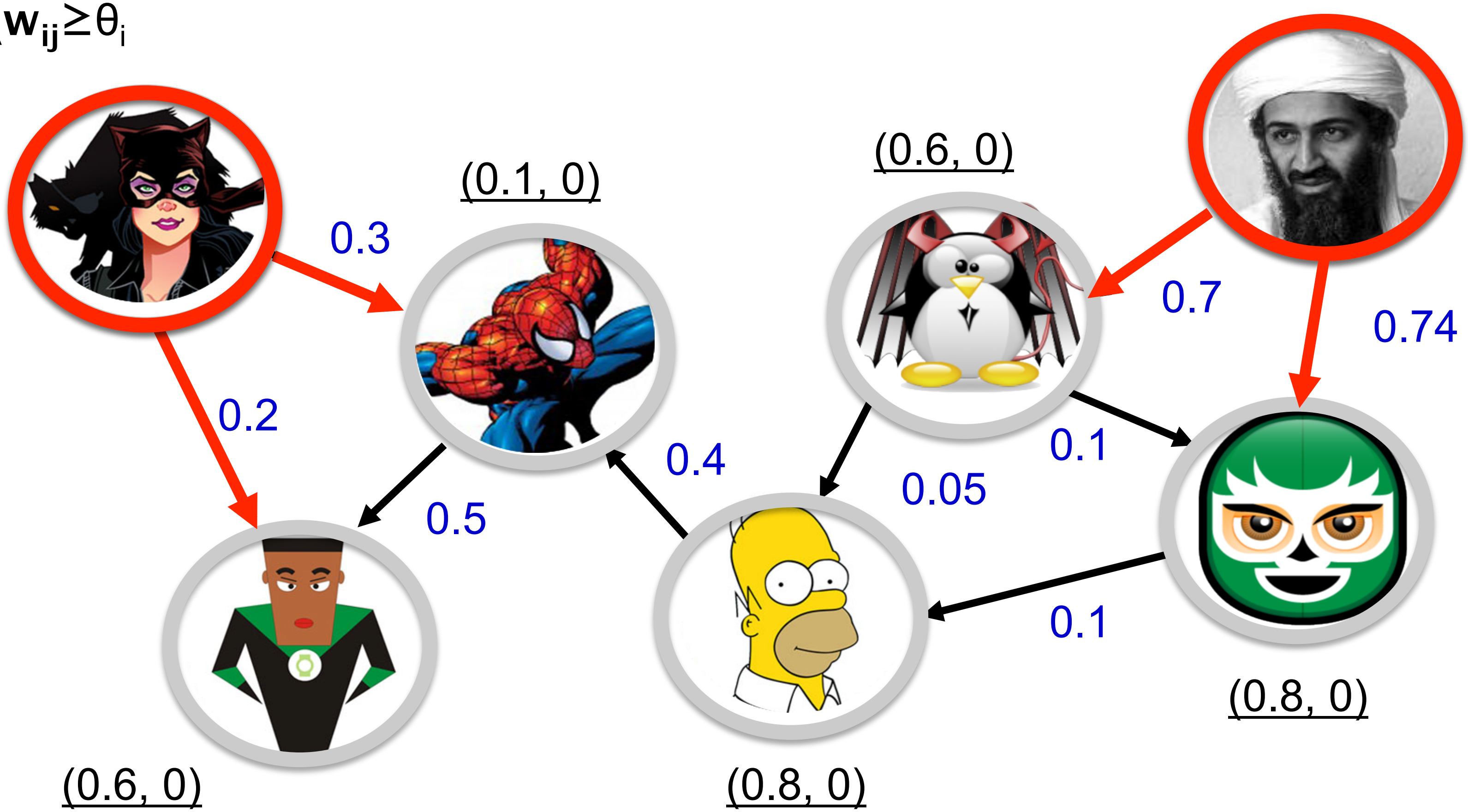
Independent Cascade Model

- Each edge is associated with a probability p_{ij}
- At first time stamp, some nodes become *active* while others are left *inactive*.
- Once a node i becomes *active*, it has a single chance to activate each of its *inactive* neighbor j with the associated probability.



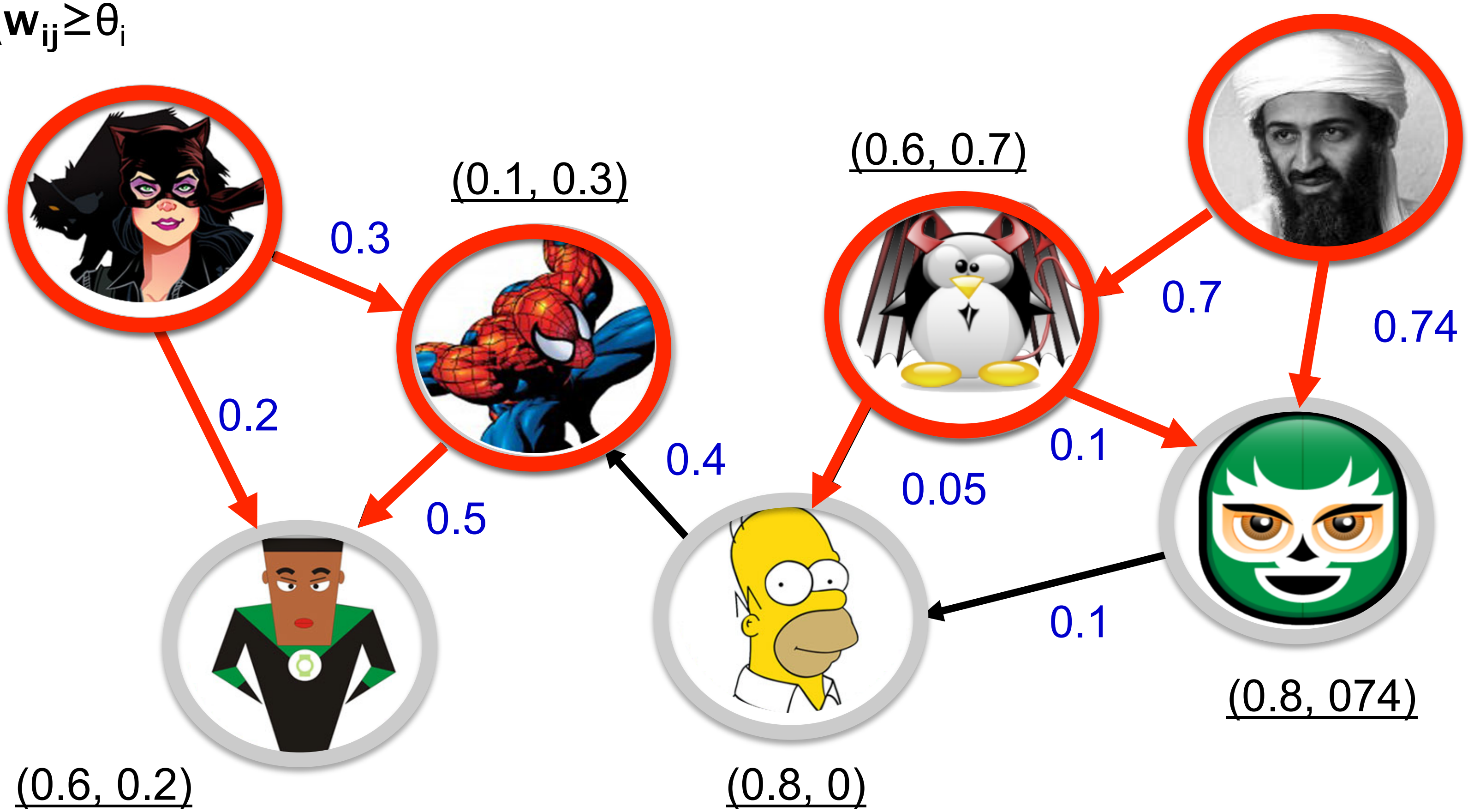
Linear Threshold Model

- Each edge is associated with a weight w_{ij}, s.t. $\boldsymbol{\Sigma w_{ij} \leq 1}$
- For each node i, assign a random threshold $\theta_i \sim U[0, 1]$
- At first time stamp, some nodes become *active* while others are left *inactive*.
- A node i becomes *active* when its weighted active neighbors exceed the threshold $\boldsymbol{\Sigma_{j \in A} w_{ij}} \geq \theta_i$



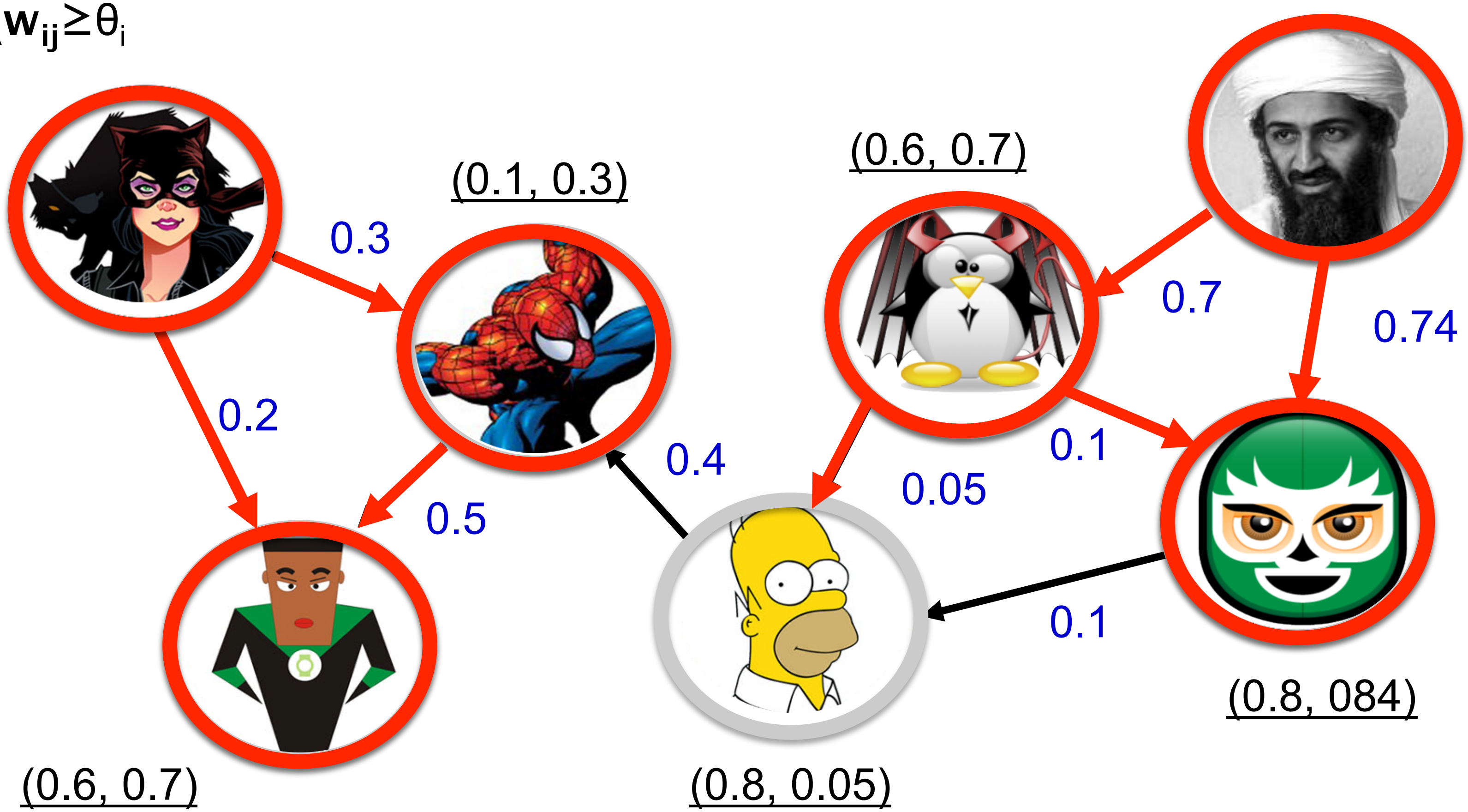
Linear Threshold Model

- Each edge is associated with a weight w_{ij}, s.t. $\mathbf{\Sigma w_{ij} \leq 1}$
- For each node i, assign a random threshold $\theta_i \sim U[0, 1]$
- At first time stamp, some nodes become *active* while others are left *inactive*.
- A node i becomes *active* when its weighted active neighbors exceed the threshold $\Sigma_{j \in A} w_{ij} \geq \theta_i$



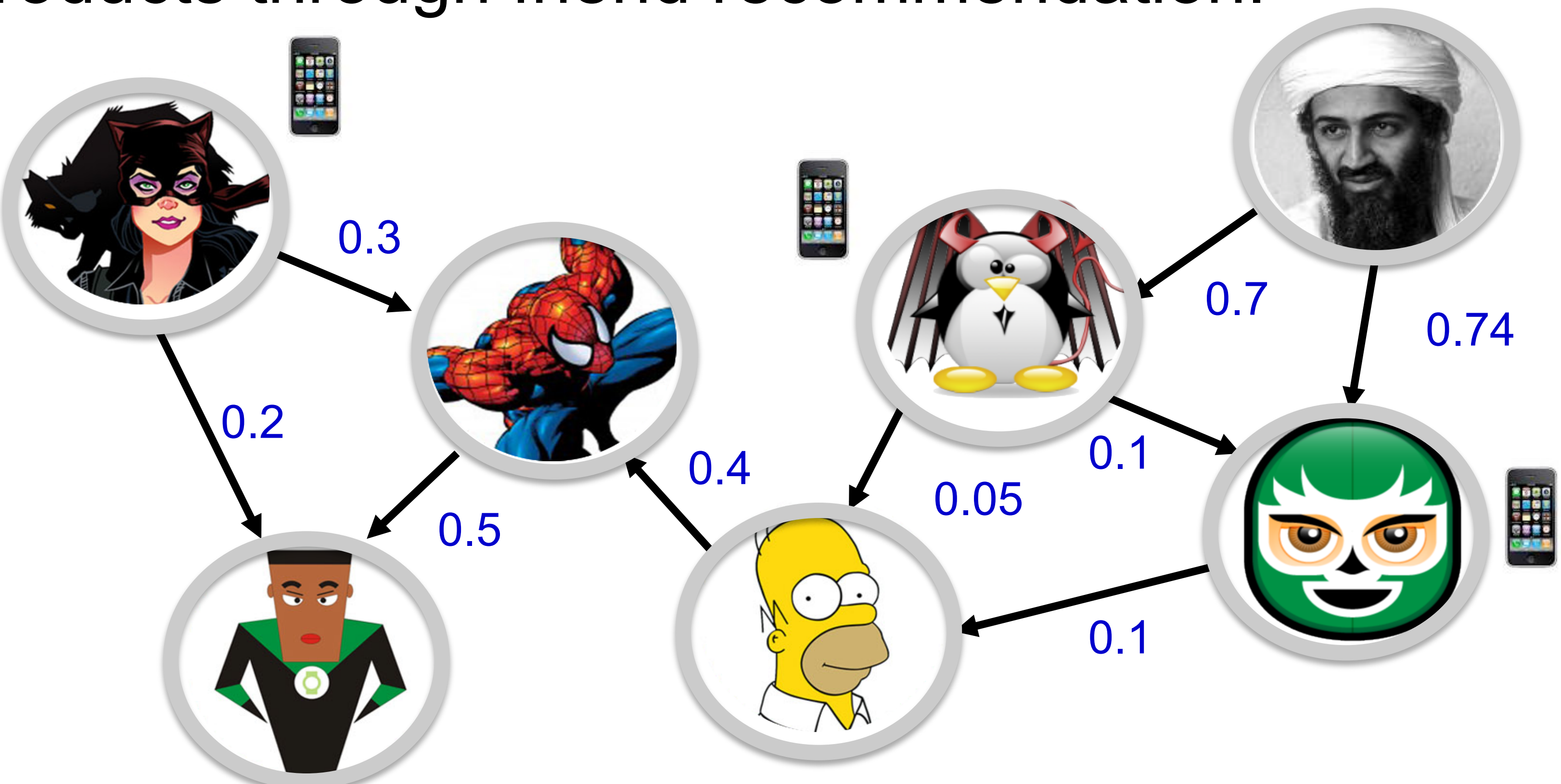
Linear Threshold Model

- Each edge is associated with a weight w_{ij}, s.t. $\boldsymbol{\Sigma w_{ij} \leq 1}$
- For each node i, assign a random threshold θ_i ~U[0, 1]
- At first time stamp, some nodes become *active* while others are left *inactive*.
- A node i becomes *active* when its weighted active neighbors exceed the threshold $\boldsymbol{\Sigma_{j \in A} w_{ij}} \geq \theta_i$



Influence Maximization

- Initially targeting a few “influential” seeds, to trigger a maximal number of individuals to adopt the opinions/ products through friend recommendation.



D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In KDD’03, pages 137–146, 2003

Influence Maximization

- Influence spread *F(S)*
  - *S* is the initial set of activated nodes, i.e., "seed set"
  - Defined as the expected number of active nodes in the end
- Objective
  - For a given budget *k*
  - Find $S^* = \arg\max F(S)$, $|S| = k$
- Challenge
  - The optimization problem is NP-hard

D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In KDD'03, pages 137–146, 2003

Greedy Algorithm

- Initialize the seed set as an empty set $S \leftarrow \varnothing$
- For *k* times, select a node i which can optimize the marginal gain:

$$i \leftarrow \arg\max[F(S \cup \{i\}) - F(S)]$$

$$S \leftarrow S \cup \{i\}$$

- A performance guarantee?
  – The solution obtained by Greedy is better than 63% (1-1/e) of the optimal solution

$$F(S) \geq (1 - \frac{1}{e})F(S^*)$$

D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In KDD'03, pages 137–146, 2003

Key Question

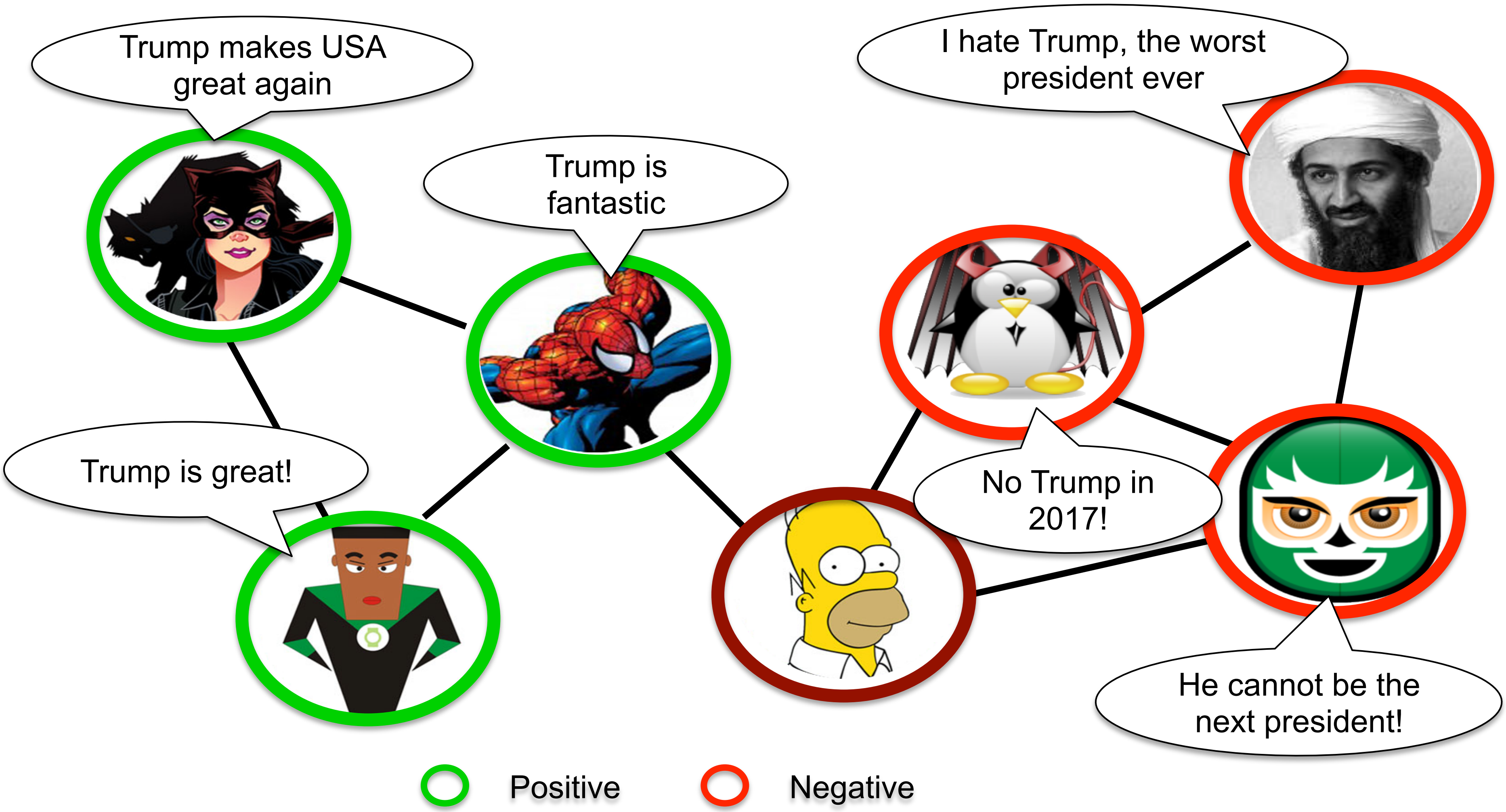
- How to obtain the weighted edges used in IC or LT models?

- How shall we learn the influence between two particular individuals?
  - Factors that affects social influence
    - Users' personal interests to a topic
    - Users' social roles

How Does Personal Interest Affect Social Influence?

Jie Tang, Jimeng Sun, Chi Wang, and Zi Yang. **Social Influence Analysis in Large-scale Networks.** KDD 2009.

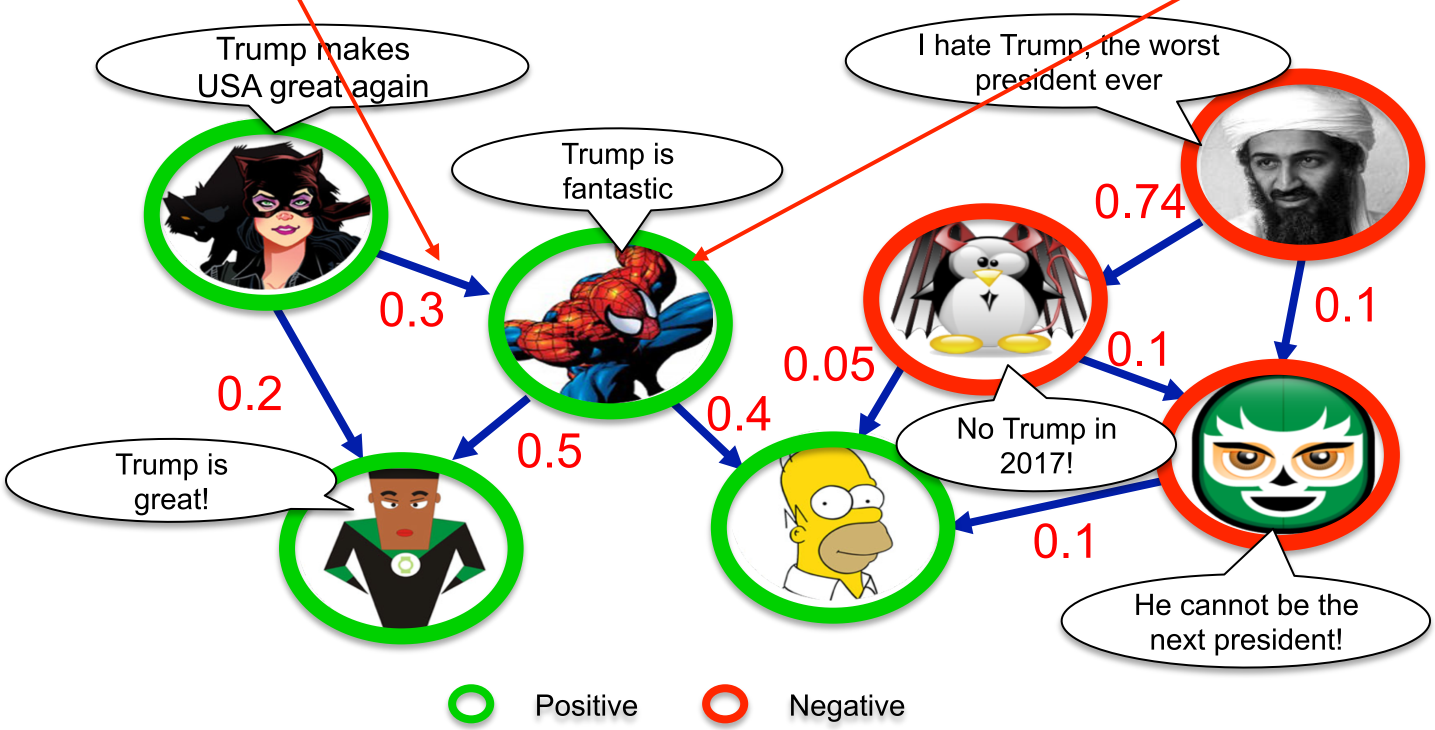
User Opinion and Influence: “Love Trump”



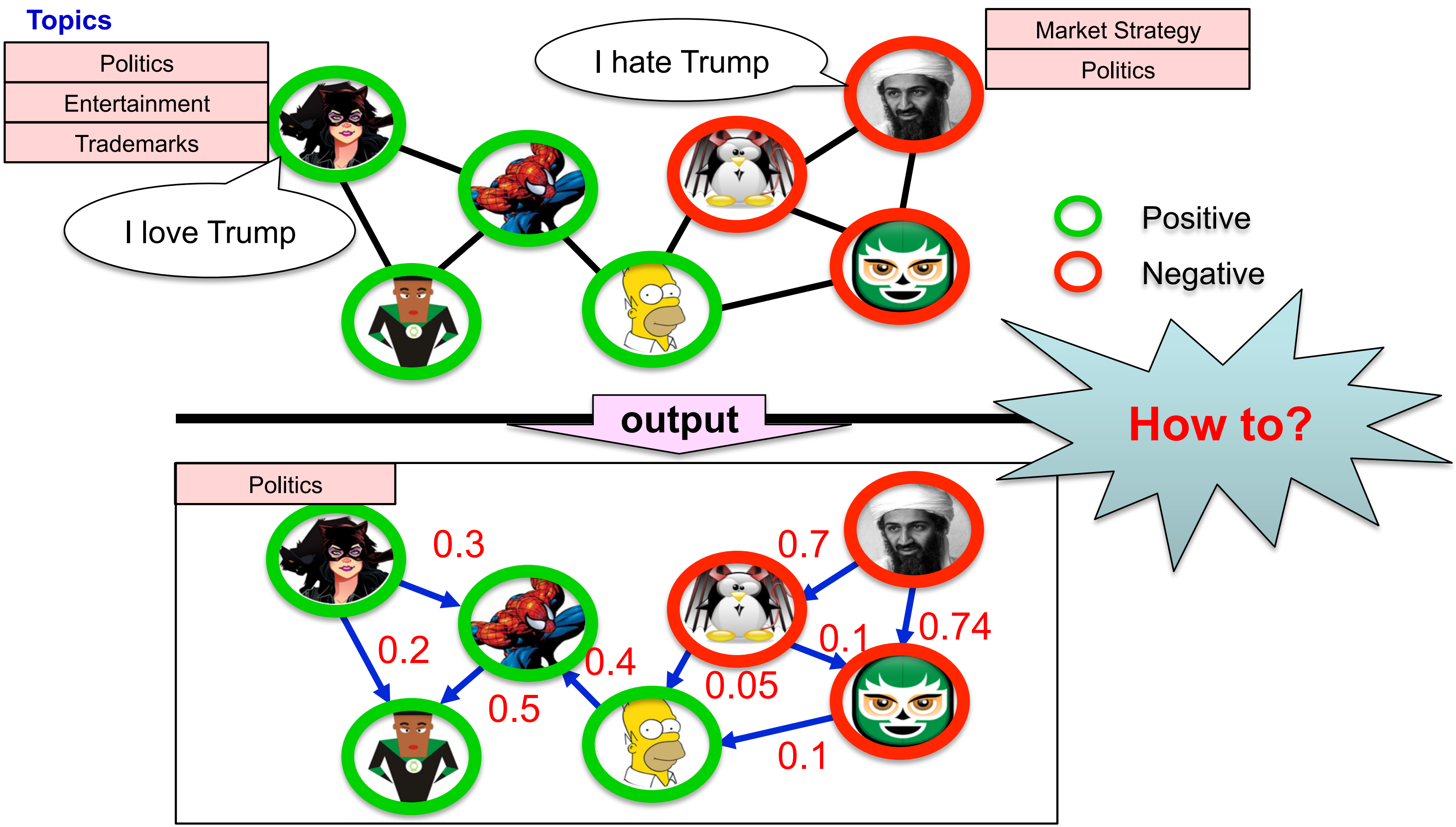
Learn Multiple Aspect Social Influence

1 Who influenced who? What is the **influence probability**?

2 How to differentiate social influences from **multiple aspects**?

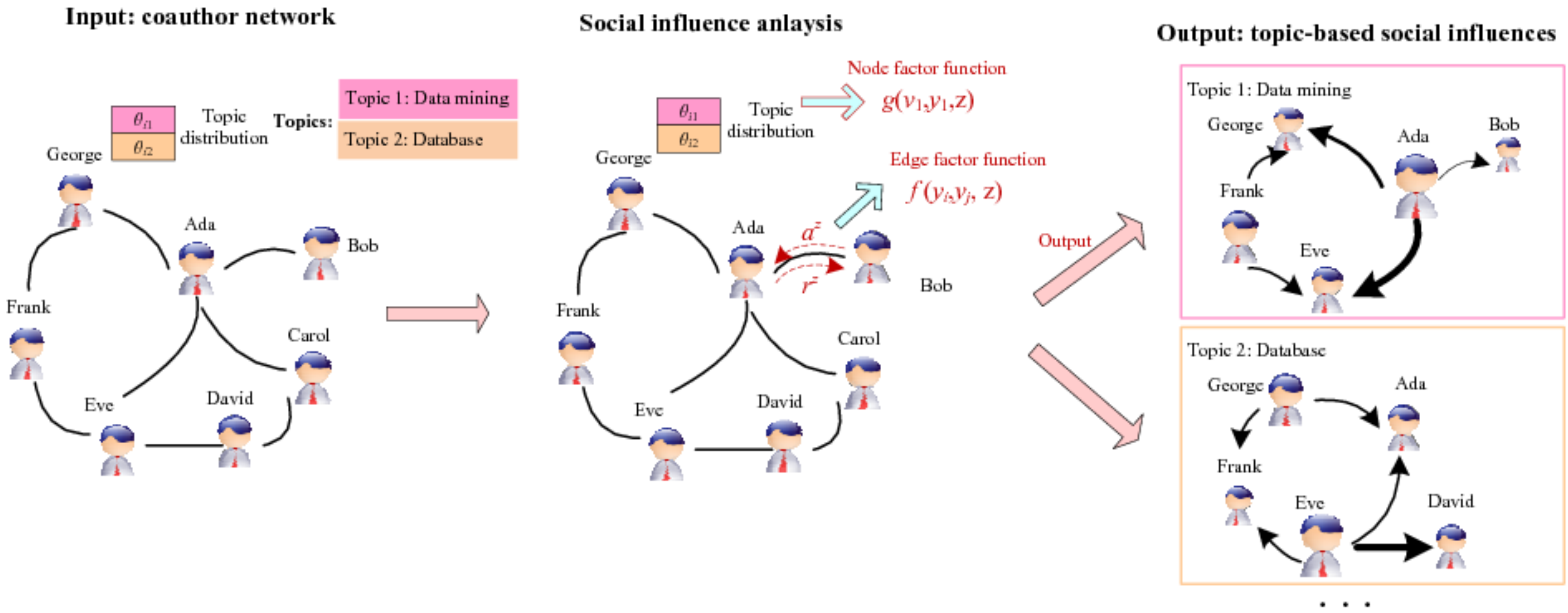


Formulation: Learning Topic-based Social Influence

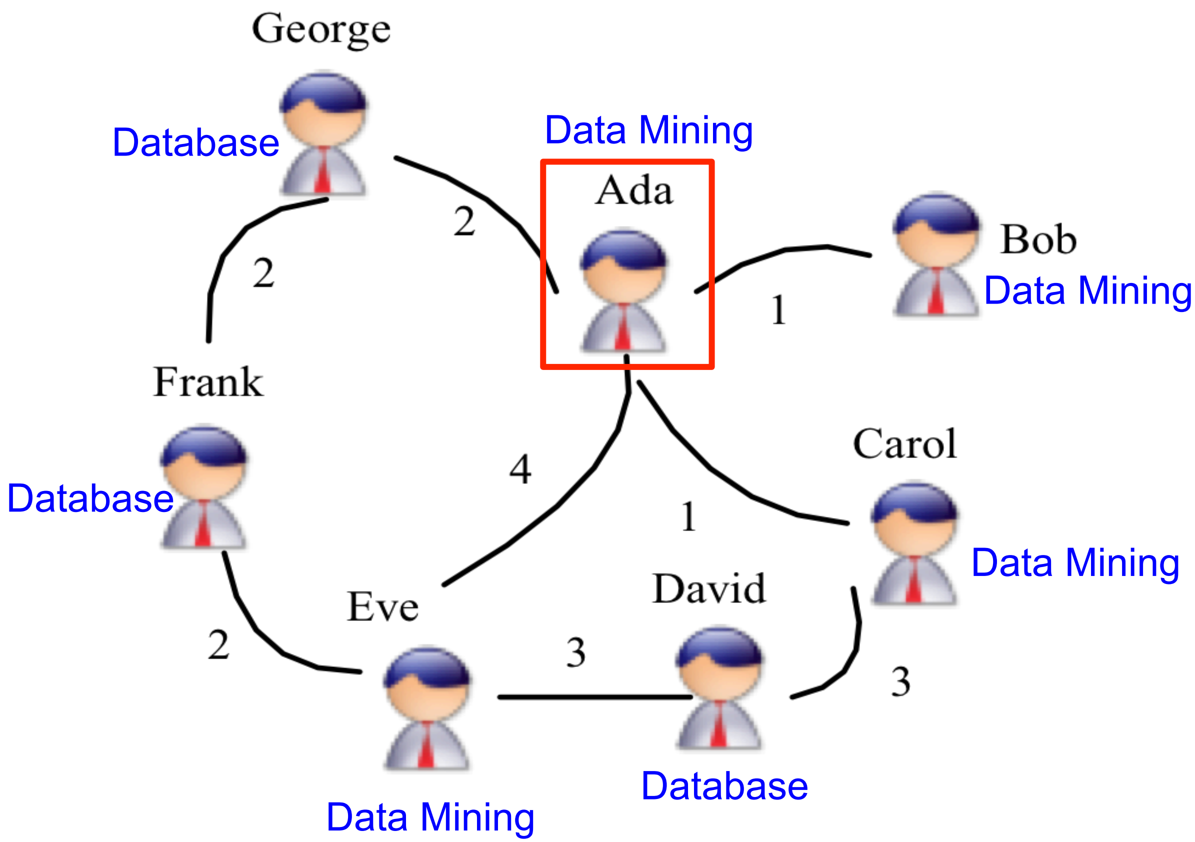


Learning Topic-based Social Influence

- Social network -> Topical influence network



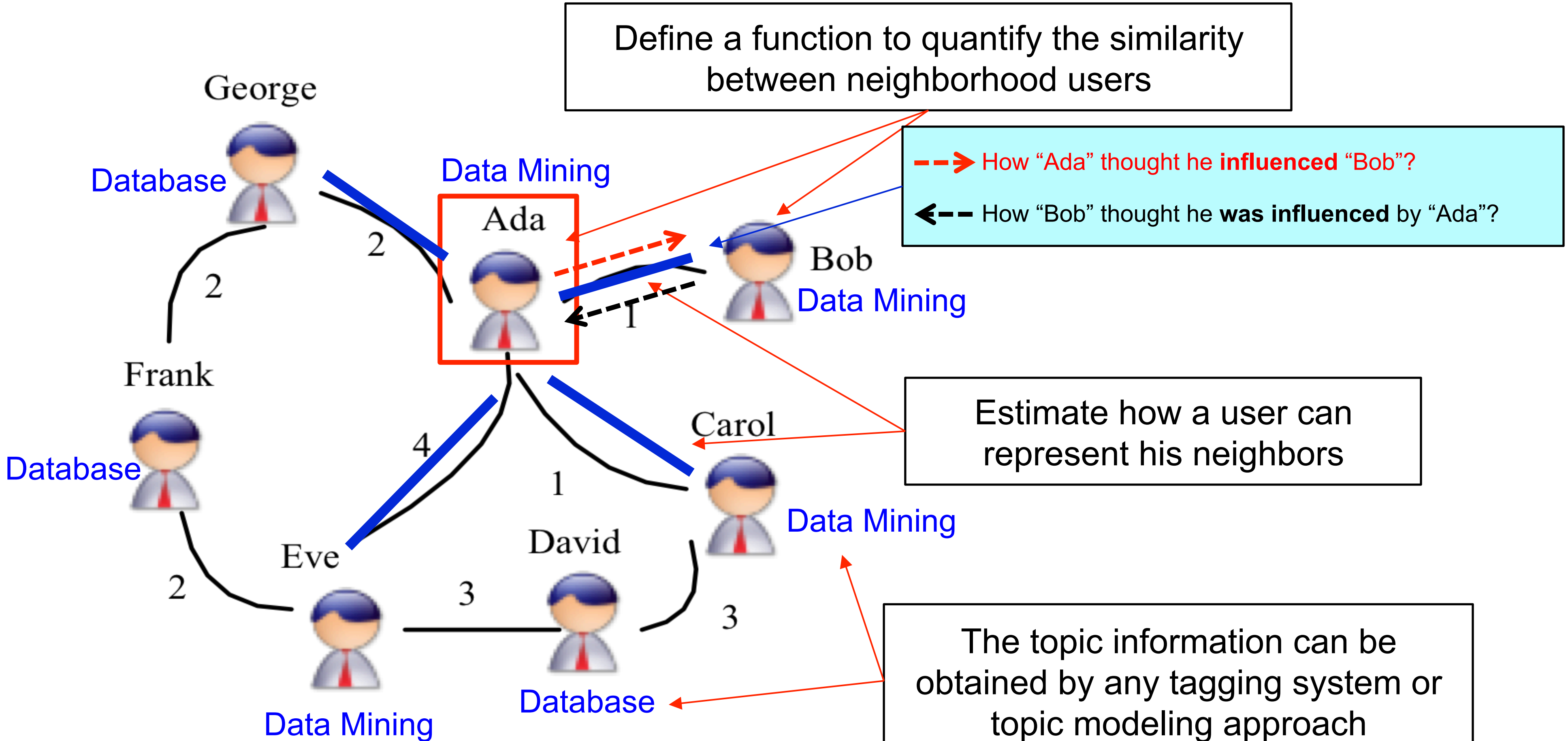
The Solution: Topical Affinity Propagation



Basic Idea:
If a user is located in the center of a community, and is "similar" to the other users, then she/he would have a strong influence on the other users.
—Homophily theory

[1] Jie Tang, Jimeng Sun, Chi Wang, and Zi Yang. **Social Influence Analysis in Large-scale Networks.** KDD 2009. (Top 10 cited paper among all papers published at KDD in the past 10 years)

The Solution: Topical Affinity Propagation



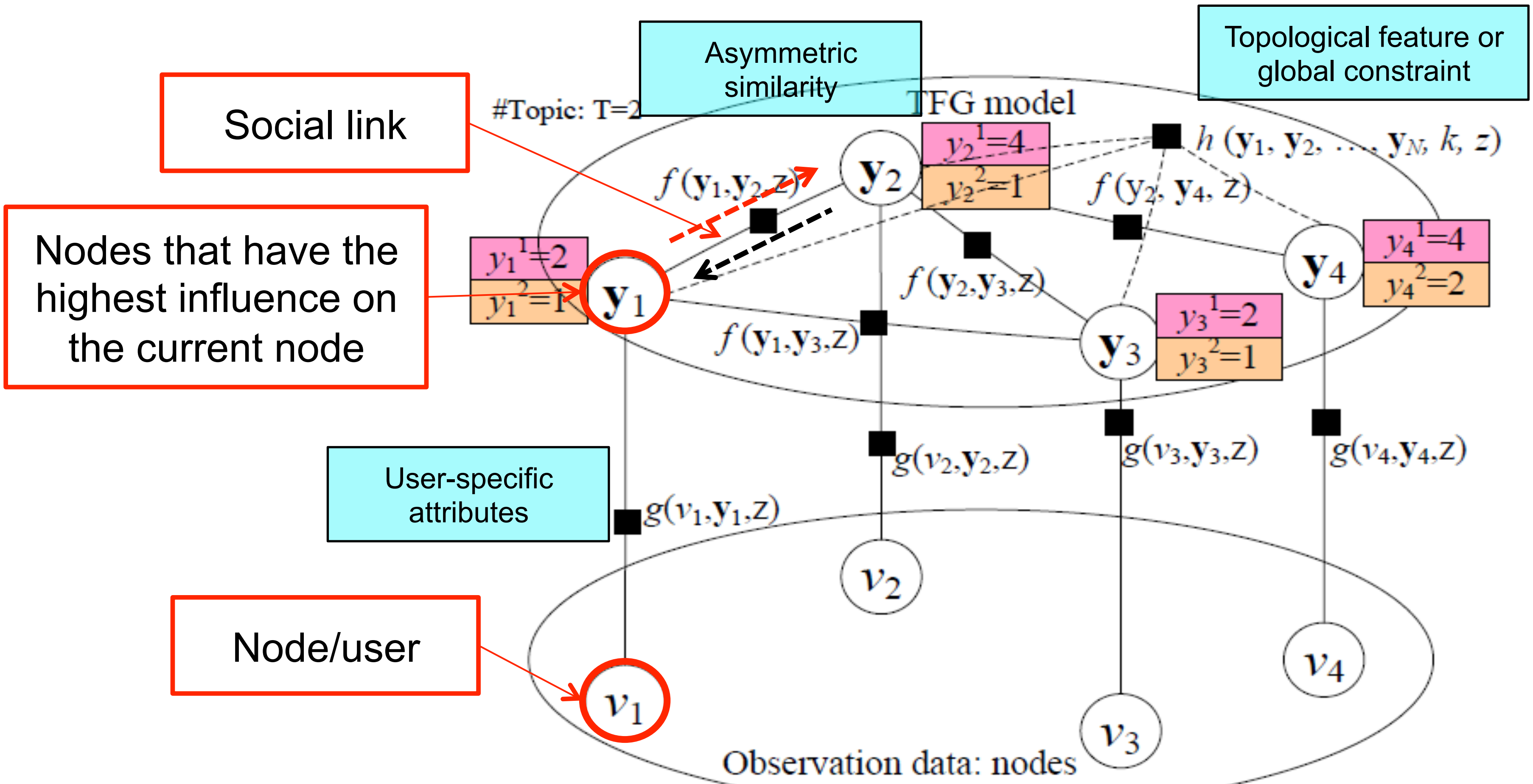
[1] Jie Tang, Jimeng Sun, Chi Wang, and Zi Yang. **Social Influence Analysis in Large-scale Networks.** KDD 2009. (Top 10 cited paper among all papers published at KDD in the past 10 years)

The Solution: Topical Affinity Propagation

- Topical Affinity Propagation
  - Topical Factor Graph model
  - Efficient learning algorithm
  - Distributed implementation

[1] Jie Tang, Jimeng Sun, Chi Wang, and Zi Yang. **Social Influence Analysis in Large-scale Networks.** KDD 2009.
(Top 10 cited paper among all papers published at KDD in the past 10 years)

Topical Factor Graph (TFG) Model



The problem is cast as identifying which node has the **highest probability** to **influence** another node on a **specific topic** along with the edge.

Topical Factor Graph (TFG)

Objective function:

$$P(\mathbf{v}, \mathbf{Y}) = \frac{1}{Z} \prod_{k=1}^{N} \prod_{z=1}^{T} \boxed{h(\mathbf{y}_1, \ldots, \mathbf{y}_N, k, z)}$$

$$\prod_{i=1}^{N} \prod_{z=1}^{T} \boxed{g(v_i, \mathbf{y}_i, z)} \prod_{e_{kl} \in E} \prod_{z=1}^{T} \boxed{f(\mathbf{y}_k, \mathbf{y}_l, z)}$$

1. How to define?

2. How to optimize?

- The learning task is to find a configuration for all $\{y_i\}$ to maximize the joint probability.

How to define (topical) feature functions?

Similarity: $w_{ij}^z = \theta_j^z \alpha_{ij}$

– Node feature function

$$g(v_i, \mathbf{y}_i, z) = \begin{cases} \dfrac{\boxed{w_{iy_i^z}^z}}{\sum_{j \in NB(i)} (w_{ij}^z + w_{ji}^z)} & y_i^z \neq i \\ \dfrac{\sum_{j \in NB(i)} w_{ji}^z}{\sum_{j \in NB(i)} (w_{ij}^z + w_{ji}^z)} & y_i^z = i \end{cases}$$

– Edge feature function

$$f(y_i, y_j) = \begin{cases} w[v_i \sim v_j] & y_i = y_j \\ 1 - w[v_i \sim v_j] & y_i \neq y_j \end{cases}$$

or simply binary

– Global feature function

$$h(\mathbf{y}_1, \ldots, \mathbf{y}_N, k, z) = \begin{cases} 0 & \text{if } y_k^z = k \text{ and } y_i^z \neq k \text{ for all } i \neq k \\ 1 & \text{otherwise.} \end{cases}$$

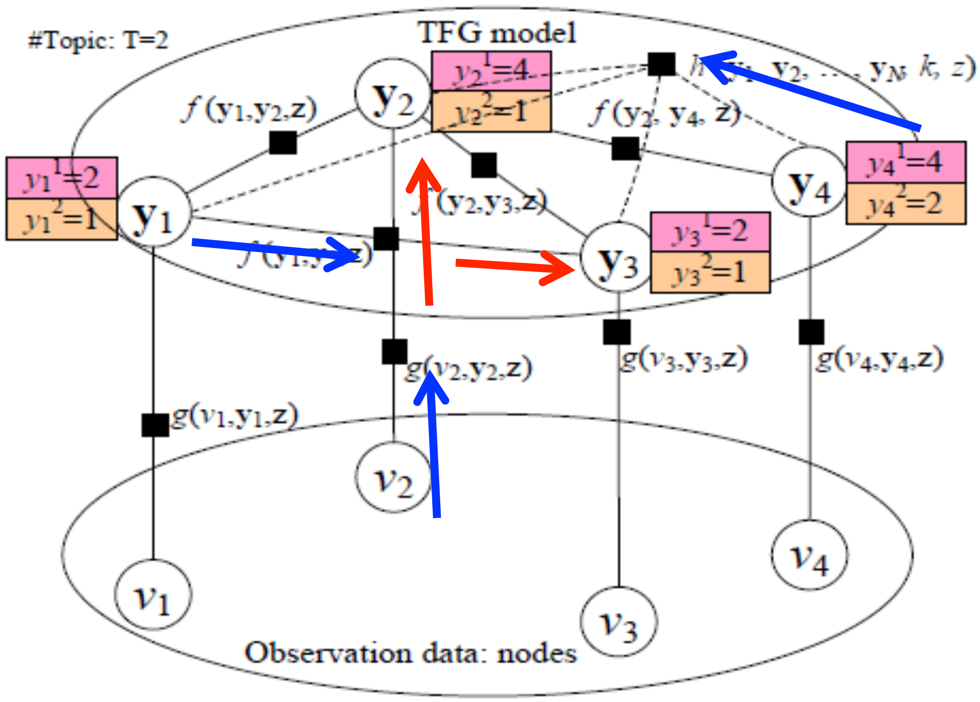
Model Learning Algorithm

Sum-product:

$$\boxed{m_{y\to f}(y,z)} = \prod_{f'\sim y\setminus f} m_{f'\to y}(y,z) \prod_{z'\neq z}\prod_{f'\sim y\setminus f} m_{f'\to y}(y,z')^{(\tau_{z'z})}$$

$$\boxed{m_{f\to y}(y,z)} = \boxed{\sum_{\sim\{y\}}}\left(f(Y,z)\prod_{y'\sim f\setminus y} m_{y'\to f}(y',z)\right) + \sum_{z'\neq z}\tau_{z'z}\sum_{\sim\{y\}}\left(f(Y,z')\prod_{y'\sim f\setminus y} m_{y'\to f}(y',z')\right) \quad (4)$$

Marginal function for y on topic z



- Low efficiency!
- Not easy for distributed learning!

New TAP Learning Algorithm

1. Introduce two new variables *r* and *a, to* replace the original message *m*.

2. Design new update rules:

How user *i* thought he **influenced** user *j*?

$$\boxed{m_{ij}}$$

$$r_{ij}^z = b_{ij}^z - \max_{k \in NB(j)} \{b_{ik}^z + a_{ik}^z\}$$

$$a_{jj}^z = \max_{k \in NB(j)} \min\{r_{kj}^z, 0\}$$

$$a_{ij}^z = \min(\max\{r_{jj}^z, 0\}, -\min\{r_{jj}^z, 0\}$$
$$- \max_{k \in NB(j)\backslash\{i\}} \min\{r_{kj}^z, 0\}), i \in NB(j)$$

How user *j* thought he **was influenced** by user *i*?

The TAP Learning Algorithm

Input: $G = (V, E)$ and topic distributions $\{\theta_v\}_{v \in V}$

Output: topic-level social influence graphs $\{G_z = (V_z, E_z)\}_{z=1}^{T}$

Calculate the node feature function $g(v_i, \mathbf{y}_i, z)$;

Calculate b_{ij}^z according to Eq. 8;

$$b_{ij}^z = \log \frac{g(v_i, \mathbf{y}_i, z)|_{y_i^z = j}}{\sum_{k \in NB(i) \cup \{i\}} g(v_i, \mathbf{y}_i, z)|_{y_i^z = k}}$$

Initialize all $\{r_{ij}^z\} \leftarrow 0$;

repeat

    foreach *edge-topic pair* (e_{ij}, z) **do**

        Update r_{ij}^z according to Eq. 5;

$$r_{ij}^z = b_{ij}^z - \max_{k \in NB(j)} \{b_{ik}^z + a_{ik}^z\}$$

    end

    foreach *node-topic pair* (v_j, z) **do**

        Update a_{jj}^z according to Eq. 6;

$$a_{jj}^z = \max_{k \in NB(j)} \min\{r_{kj}^z, 0\}$$

    end

    foreach *edge-topic pair* (e_{ij}, z) **do**

        Update a_{ij}^z according to Eq. 7;

$$a_{ij}^z = \min(\max\{r_{jj}^z, 0\}, -\min\{r_{jj}^z, 0\} - \max_{k \in NB(j) \setminus \{i\}} \min\{r_{kj}^z, 0\}), i \in NB(j)$$

    end

until *convergence*;

foreach *node* v_t **do**

    foreach *neighboring node* $s \in NB(t) \cup \{t\}$ **do**

        Compute μ_{st}^z according to Eq. 9;

$$\mu_{st}^z = \frac{1}{1 + e^{-(r_{ts}^z + a_{ts}^z)}}$$

    end

end

Generate $G_z = (V_z, E_z)$ for every topic z according to $\{\mu_{st}^z\}$;

Distributed TAP Learning

- Map-Reduce
  - Map: (key, value) pairs
    - $e_{ij}/a_{ij} \rightarrow e_{i*}/a_{ij}$; $e_{ij}/b_{ij} \rightarrow e_{i*}/b_{ij}$; $e_{ij}/r_{ij} \rightarrow e_{*j}/r_{ij}$.
  - Reduce: (key, value) pairs
    - e_{ij} / * $\rightarrow$ new r_{ij}; e_{ij}/* $\rightarrow$ new a_{ij}
- For the global feature function

THEOREM 1. *If the global feature function h can be factorized into $h = \prod_{k=1}^{N} h_k$, for every $i \in \{1, \ldots, N\}, y_i \neq k, y_i' \neq k, h_k(y_1, \ldots, y_i, \ldots, y_N) = h_k(y_1, \ldots, y_i', \ldots, y_N)$, then the message passing update rules can be simplified to influence update rules.* ■

Experiment

- Data set: (ArnetMiner.org and Wikipedia)
  - **Coauthor** dataset： 640,134 authors and 1,554,643 coauthor relations
  - **Citation** dataset: 2,329,760 papers and 12,710,347 citations between these papers
  - **Film** dataset: 18,518 films, 7,211 directors, 10,128 actors, and 9,784 writers
- Evaluation measures
  - Case study
  - CPU time
  - Application

Influential nodes on different topics

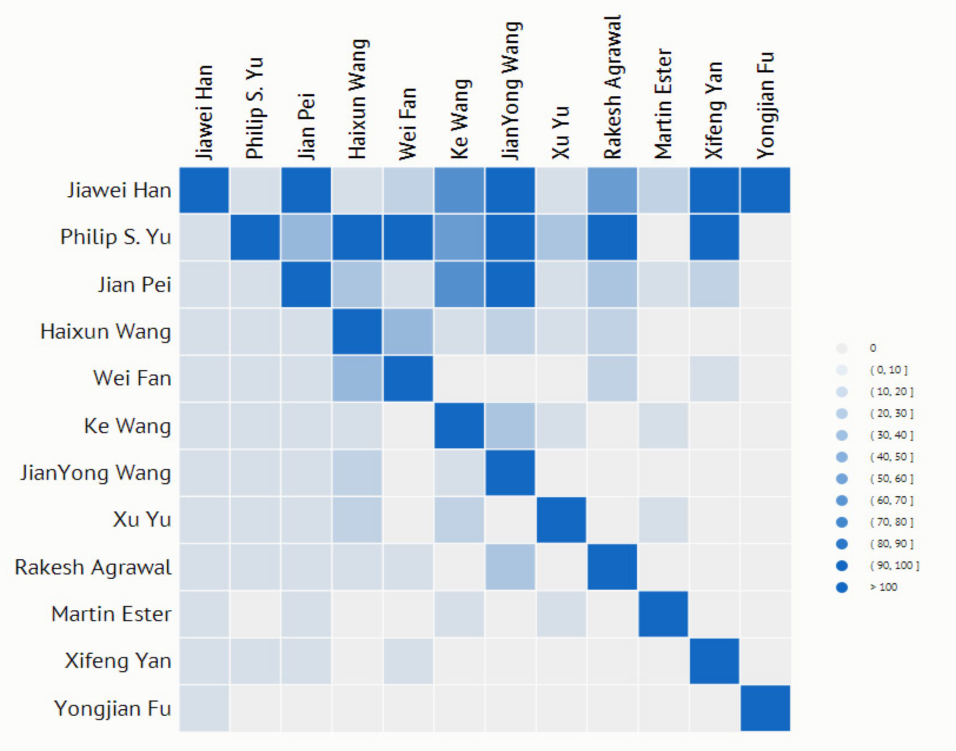
Dataset	Topic	Representative Nodes
Author	Data Mining	Heikki Mannila, Philip S. Yu, Dimitrios Gunopulos, Jiawei Han, Christos Faloutsos, Bing Liu, Vipin Kumar, Tom M. Mitchell, Wei Wang, Qiang Yang, Xindong Wu, Jeffrey Xu Yu, Osmar R. Zaiane
	Machine Learning	Pat Langley, Alex Waibel, Trevor Darrell, C. Lee Giles, Terrence J. Sejnowski, Samy Bengio, Daphne Koller, Luc De Raedt, Vasant Honavar, Floriana Esposito, Bernhard Scholkopf
	Database System	Gerhard Weikum, John Mylopoulos, Michael Stonebraker, Barbara Pernici, Philip S. Yu, Sharad Mehrotra, Wei Sun, V. S. Subrahmanian, Alejandro P. Buchmann, Kian-Lee Tan, Jiawei Han
	Information Retrieval	Gerard Salton, W. Bruce Croft, Ricardo A. Baeza-Yates, James Allan, Yi Zhang, Mounia Lalmas, Zheng Chen, Ophir Frieder, Alan F. Smeaton, Rong Jin
	Web Services	Yan Wang, Liang-jie Zhang, Schahram Dustdar, Jian Yang, Fabio Casati, Wei Xu, Zakaria Maamar, Ying Li, Xin Zhang, Boualem Benatallah, Boualem Benatallah
	Semantic Web	Wolfgang Nejdl, Daniel Schwabe, Steffen Staab, Mark A. Musen, Andrew Tomkins, Juliana Freire, Carole A. Goble, James A. Hendler, Rudi Studer, Enrico Motta
	Bayesian Network	Daphne Koller, Paul R. Cohen, Floriana Esposito, Henri Prade, Michael I. Jordan, Didier Dubois, David Heckerman, Philippe Smets
Citation	Data Mining	Fast Algorithms for Mining Association Rules in Large Databases, Using Segmented Right-Deep Trees for the Execution of Pipelined Hash Joins, Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data, Discovery of Multiple-Level Association Rules from Large Databases, Interleaving a Join Sequence with Semijoins in Distributed Query Processing
	Machine Learning	Object Recognition with Gradient-Based Learning, Correctness of Local Probability Propagation in Graphical Models with Loops, A Learning Theorem for Networks at Detailed Stochastic Equilibrium, The Power of Amnesia: Learning Probabilistic Automata with Variable Memory Length, A Unifying Review of Linear Gaussian Models
	Database System	Mediators in the Architecture of Future Information Systems, Database Techniques for the World-Wide Web: A Survey, The R*-Tree: An Efficient and Robust Access Method for Points and Rectangles, Fast Algorithms for Mining Association Rules in Large Databases
	Web Services	The Web Service Modeling Framework WSMF, Interval Timed Coloured Petri Nets and their Analysis, The design and implementation of real-time schedulers in RED-linux, The Self-Serv Environment for Web Services Composition
	Web Mining	Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data, Fast Algorithms for Mining Association Rules in Large Databases, The OO-Binary Relationship Model: A Truly Object Oriented Conceptual Model, Distributions of Surfers' Paths Through the World Wide Web: Empirical Characterizations, Improving Fault Tolerance and Supporting Partial Writes in Structured Coterie Protocols for Replicated Objects
	Semantic Web	FaCT and iFaCT, The GRAIL concept modelling language for medical terminology, Semantic Integration of Semistructured and Structured Data Sources, Description of the RACER System and its Applications, DL-Lite: Practical Reasoning for Rich Dls

Social Influence Sub-graph on "Data mining"

Table 4: Dynamic influence analysis for Dr. Jian Pei during 2000-2009. Due to space limitation, we only list coauthors who most influence on/by Dr. Pei in each time window.

Year	Pairwise	Influence
2000 - 2001	Influence on Dr. Pei	Jiawei Han (0.4961)
	Influenced by Dr. Pei	Jiawei Han (0.0082)
2002 - 2003	Influence on Dr. Pei	Jiawei Han (0.4045), Ke Wang (0.0418), Jianyong Wang (0.019), Xifeng Yan (0.007), Shiwei Tang (0.0052)
	Influenced by Dr. Pei	Shiwei Tang (0.436), Hasan M.Jamil (0.4289), Xifeng Yan (0.2192), Jianyong Wang (0.1667), Ke Wang (0.0687)
2004 - 2005	Influence on Dr. Pei	Jiawei Han (0.2364), Ke Wang (0.0328), Wei Wang (0.0294), Jianyong Wang (0.0248), Philip S. Yu (0.0156)
	Influenced by Dr. Pei	Chun Tang (0.5929), Shiwei Tang (0.5426), Hasan M.Jamil (0.3318), Jianyong Wang (0.1609), Xifeng Yan (0.1458), Yan Huang (0.1054)
2006 - 2007	Influence on Dr. Pei	Jiawei Han (0.1201), Ke Wang (0.0351), Wei Wang (0.0226), Jianyong Wang (0.018), Ada Wai-Chee Fu (0.0125)
	Influenced by Jian Pei	Chun Tang (0.6095), Shiwei Tang (0.6067), Byung-Won On (0.4599), Hasan M.Jamil (0.3433), Jaewoo Kang (0.3386)
2008 - 2009	Influence on Dr. Pei	Jiawei Han (0.2202), Ke Wang (0.0234), Ada Wai-Chee Fu (0.0208), Wei Wang (0.011), Jianyong Wang (0.0095)
	Influenced by Dr. Pei	ZhaoHui Tang (0.654), Chun Tang (0.6494), Shiwei Tang (0.5923), Zhengzheng Xing (0.5549), Hasan M.Jamil (0.3333), Jaewoo Kang (0.3057)

On "Data Mining" in 2009



Scalability Performance

Table 2: Scalability performance of different methods on real data sets. >10hr means that the algorithm did not terminate when the algorithm runs more than 10 hours.

Methods	Citation	Coauthor	Film
Sum-Product	N/A	>10hr	1.8 hr
Basic TAP Learning	>10hr	369s	**57s**
Distributed TAP Learning	**39.33m**	**104s**	148s

Application—Expert Finding

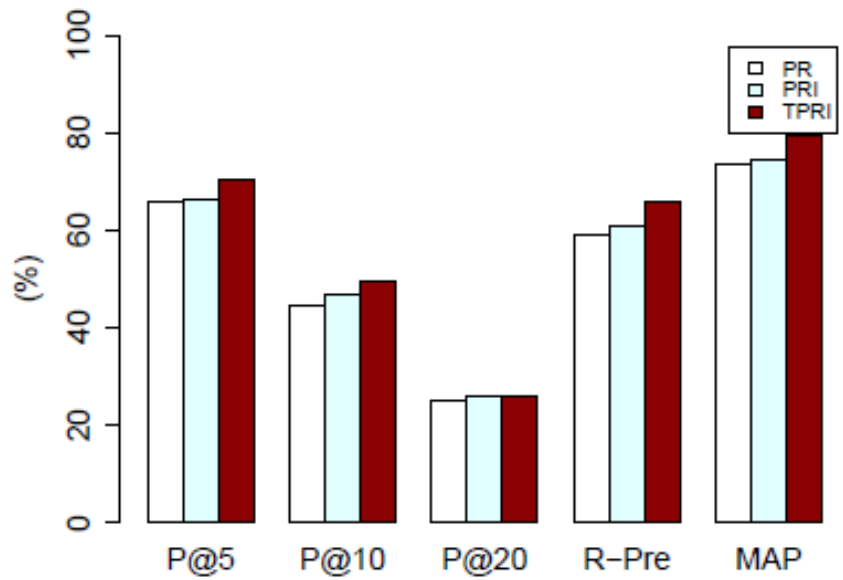


Table 7: Performance of expert finding with different approaches.

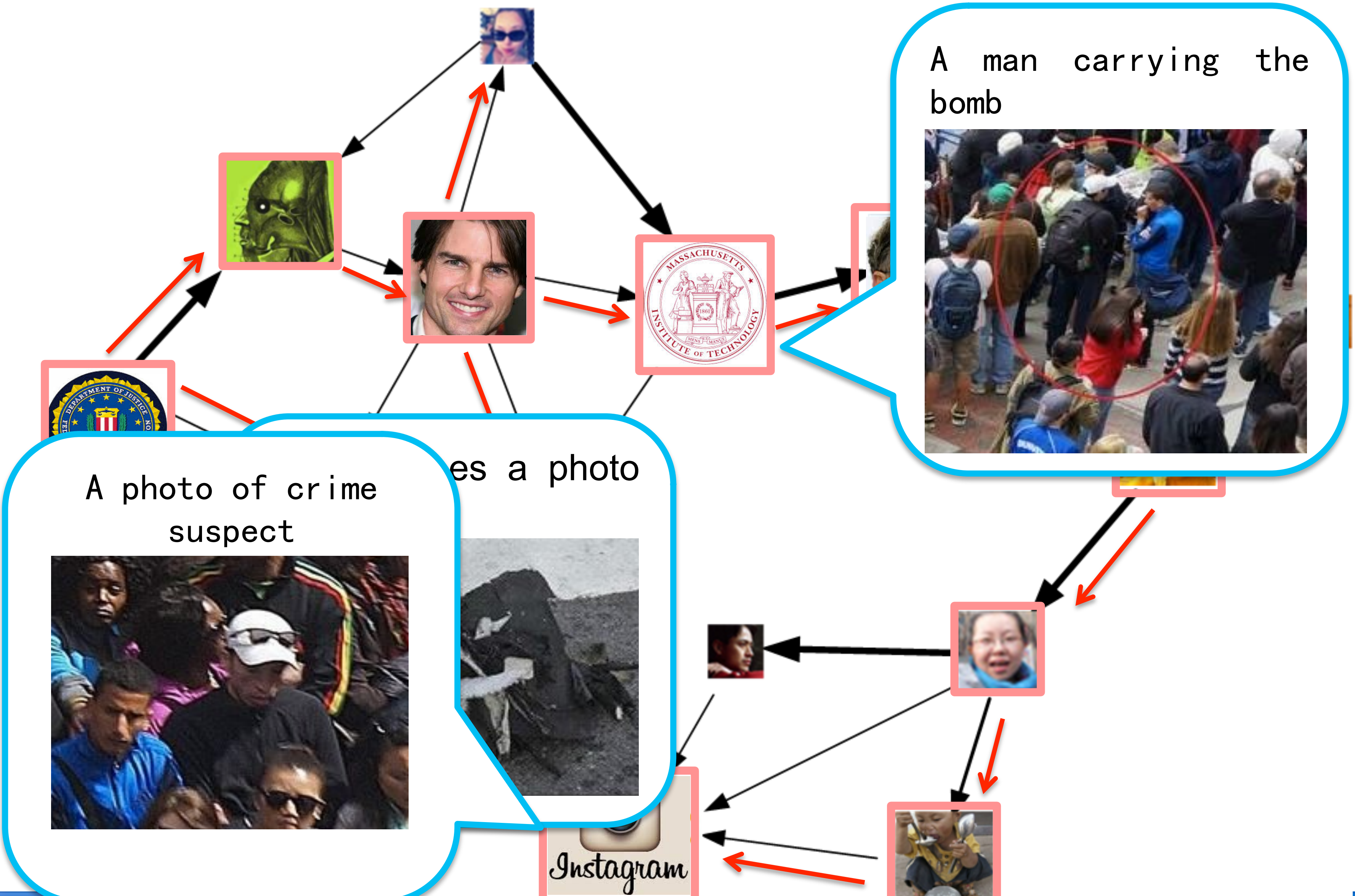
Expert finding data from (Tang, KDD08; ICDM08)

http://arnetminer.org/lab-datasets/expertfinding/

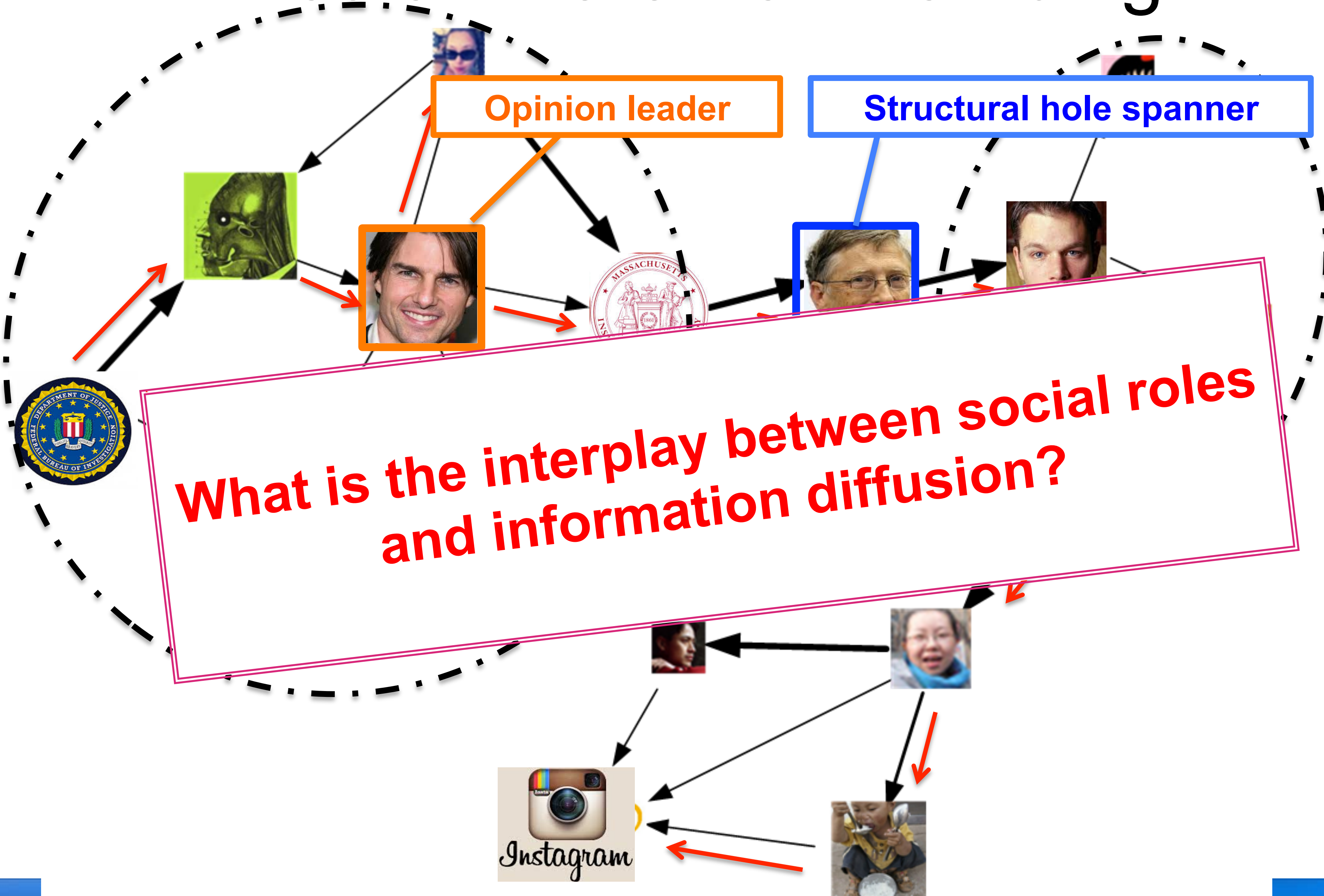
Information Diffusion

- Information diffusion, also known as diffusion of innovations, is the study of how information propagates in or between networks.

Boston Marathon Bombing



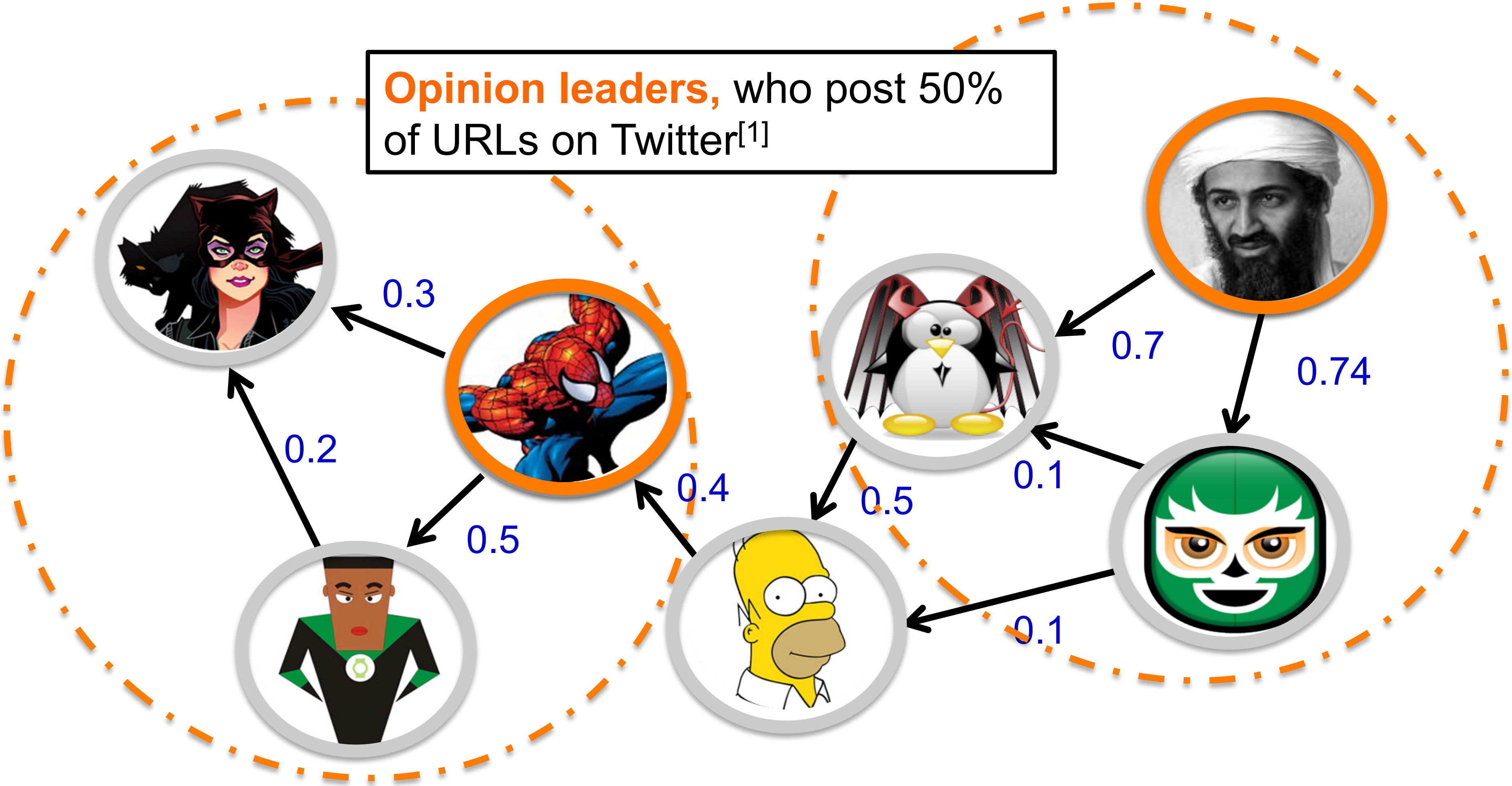
Boston Marathon Bombing



Social-Role aware Information Diffusion

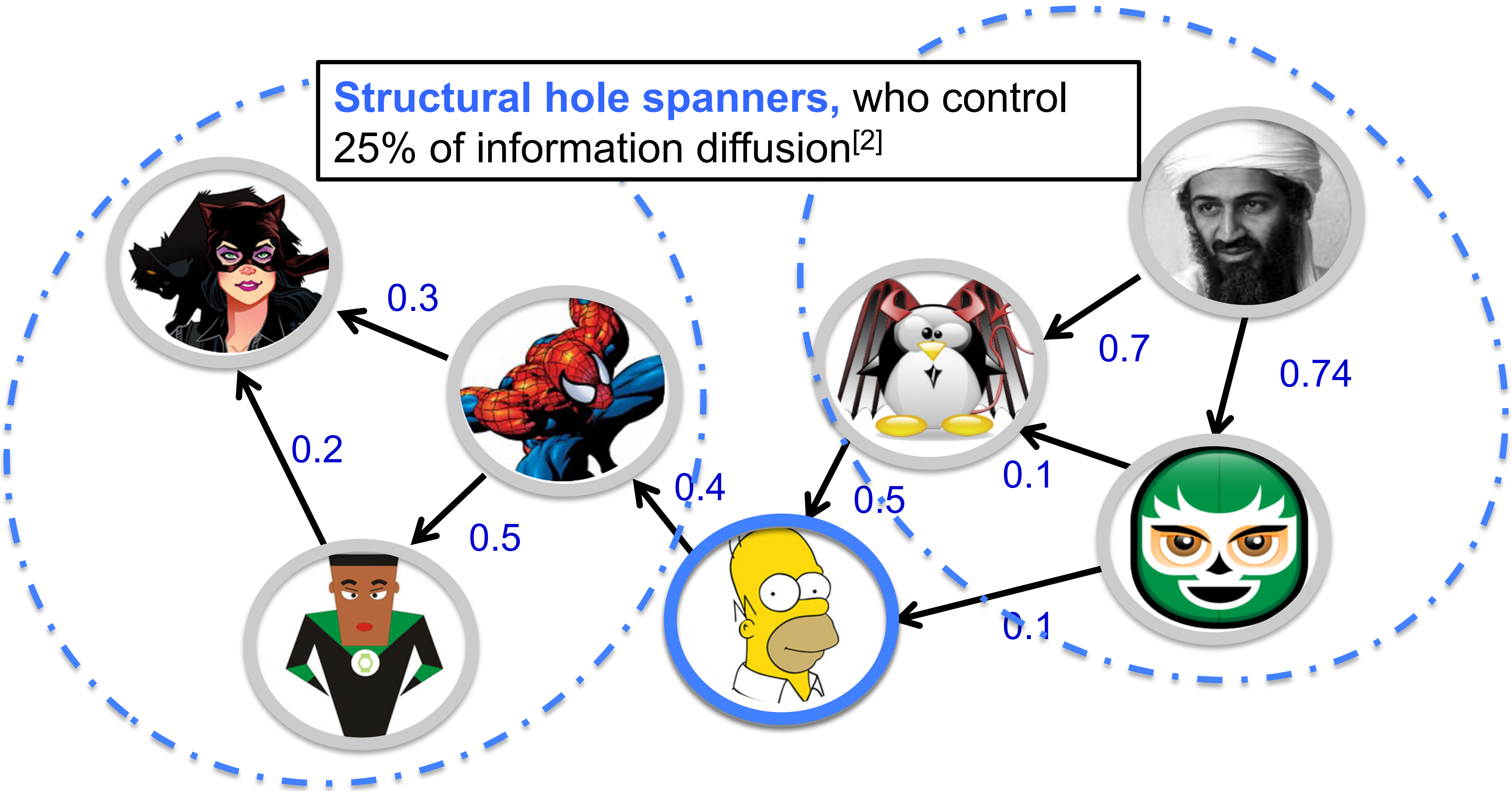
Yang Yang, Jie Tang, Cane Wing-Ki Leung, Yizhou Sun, Qicong Chen, Juanzi Li, and Qiang Yang. **RAIN: Social Role-Aware Information Diffusion.** AAAI'15, 2015.

Social Roles



[1] Wu, S.; Hofman, J. M.; Mason,W. A.; andWatts, D. J. 2011. Who says what to whom on twitter. In WWW'11 , 705–714.

Social Role

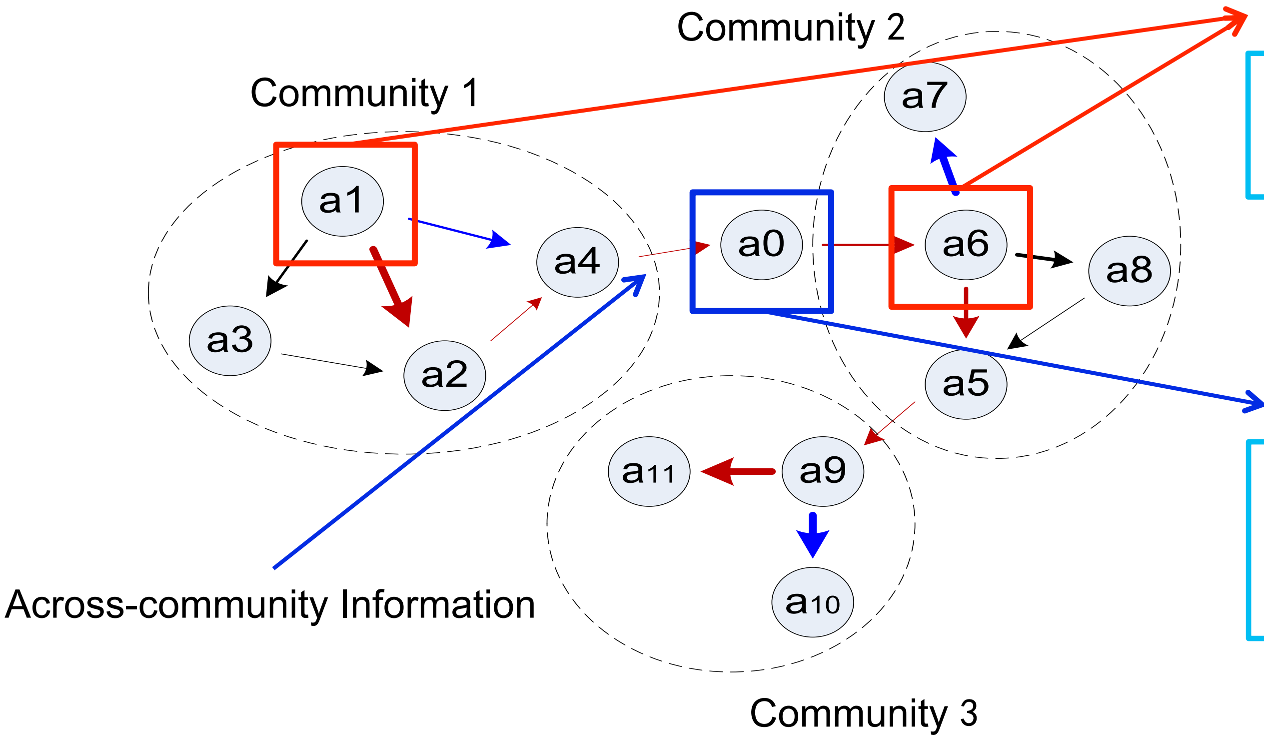


[2] Lou, T., and Tang, J. 2013. Mining structural hole spanners through information diffusion in social networks. In WWW'13 , 825–836

Social Roles

腾讯微博

>0.16 billion users
>0.17 billion posts
Complete data sets during Oct. 1st – Oct. 7th, 2012.



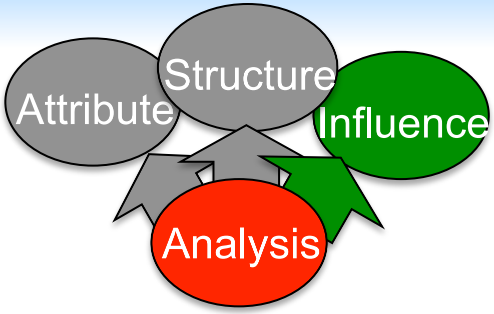
Opinion Leader

1% of users with largest PageRank scores

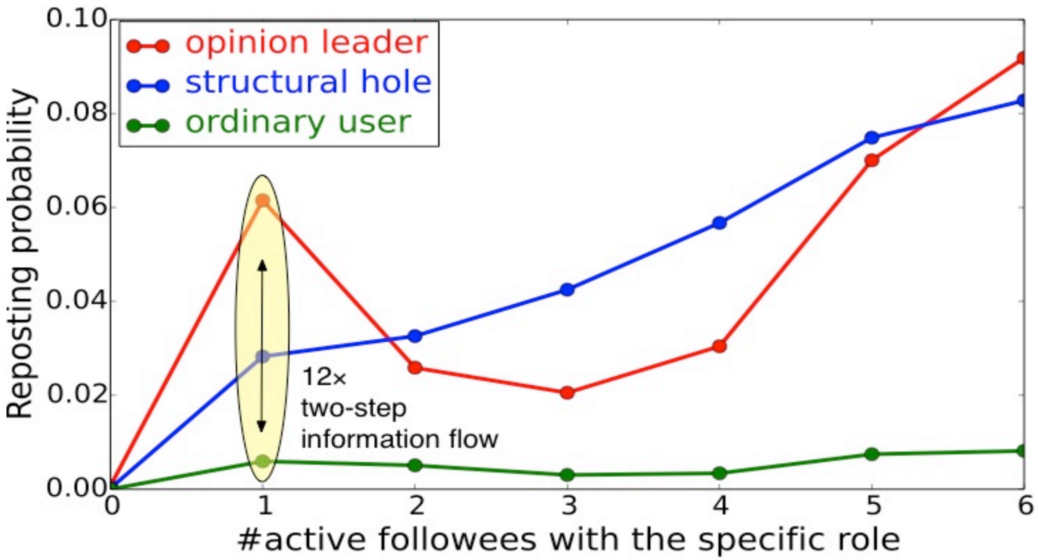
Structural Hole

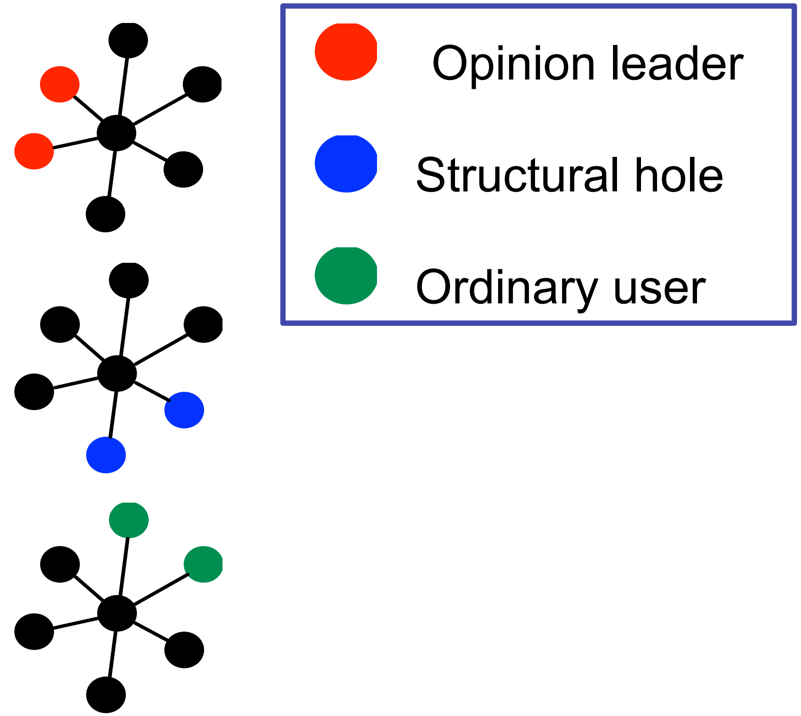
1% of users with smallest network constraint scores.

[1] S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts. Who says what to whom on twitter. In **WWW'11**, pages 705–714, 2011.
[2] T. Lou and J. Tang. Mining Structural Hole Spanners Through Information Diffusion in Social Networks. In **WWW'13**. pp. 837-848.



Influence Strength

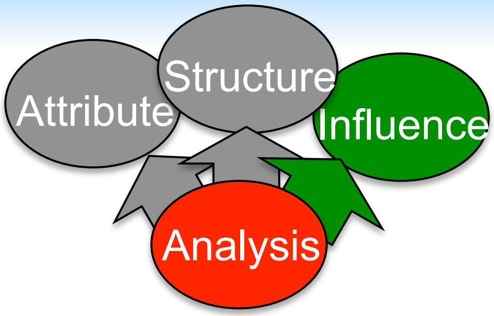




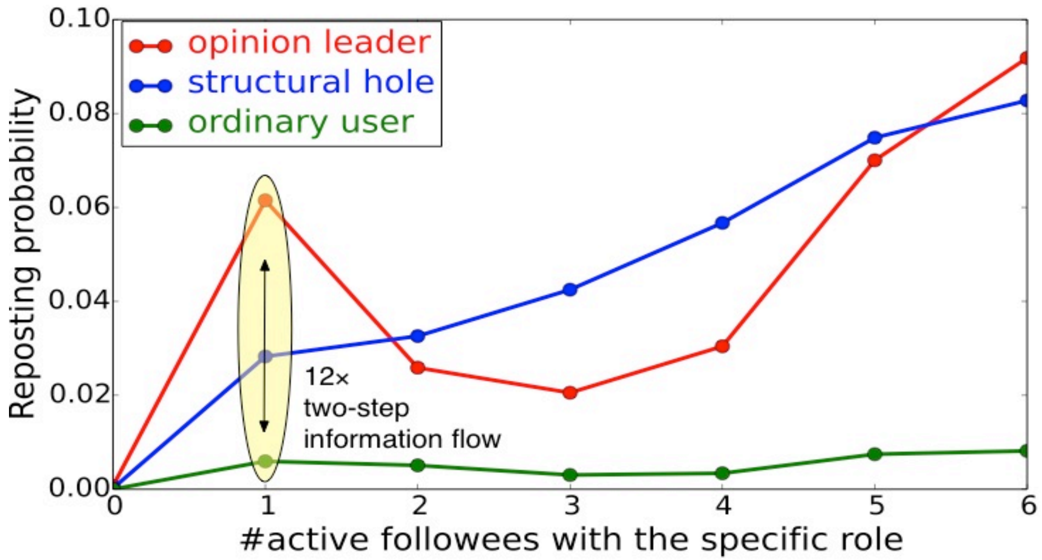
Opinion leader:

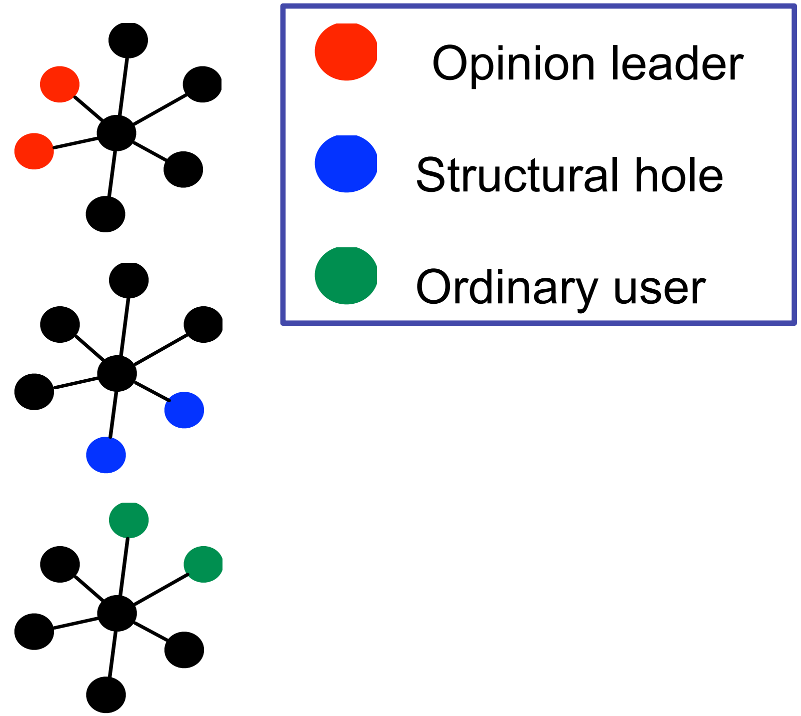
- Stage 1 - activation probability is12 times higher than ordinary user
- Stage 2 - information overload[1]: 2-3 opinion leaders are sufficient to spread a piece of information throughout a community.
- Stage 3 - information everywhere: spreading the information becomes a social norm to adopt.

[1] Lazarsfeld, P. F.; Berelson, B.; and Gaudet, H. 1944. The peoples choice: How the voter makes up his mind in a presidential election. New York: Duell, Sloan and Pearce .



Influence Strength

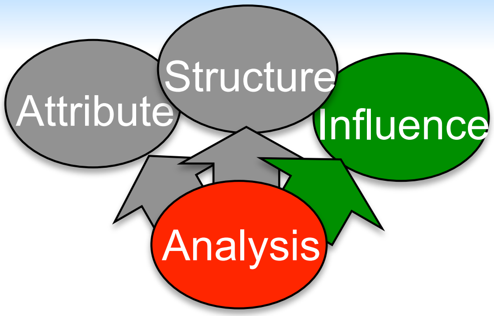




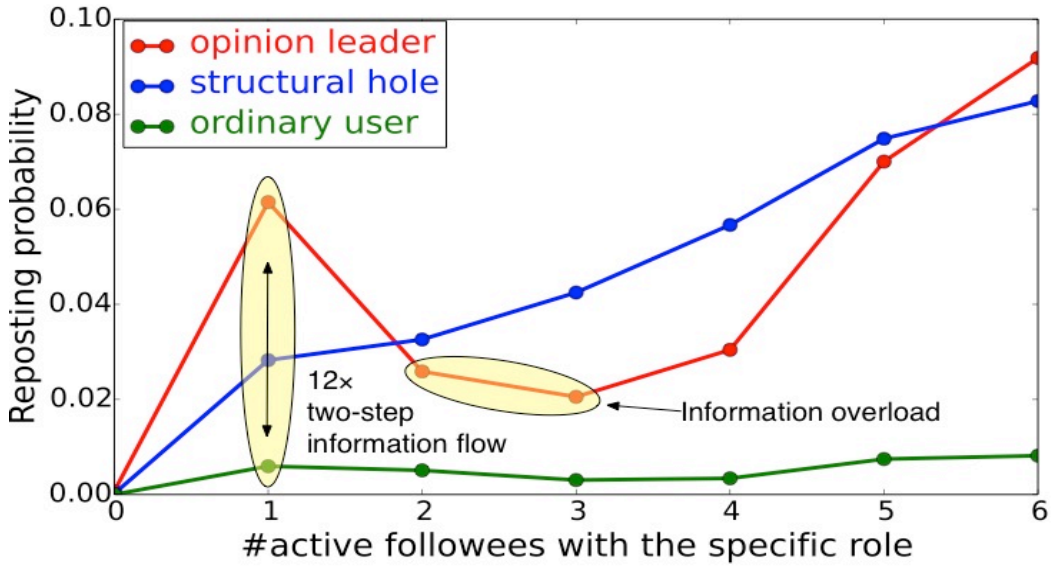
Opinion leader:

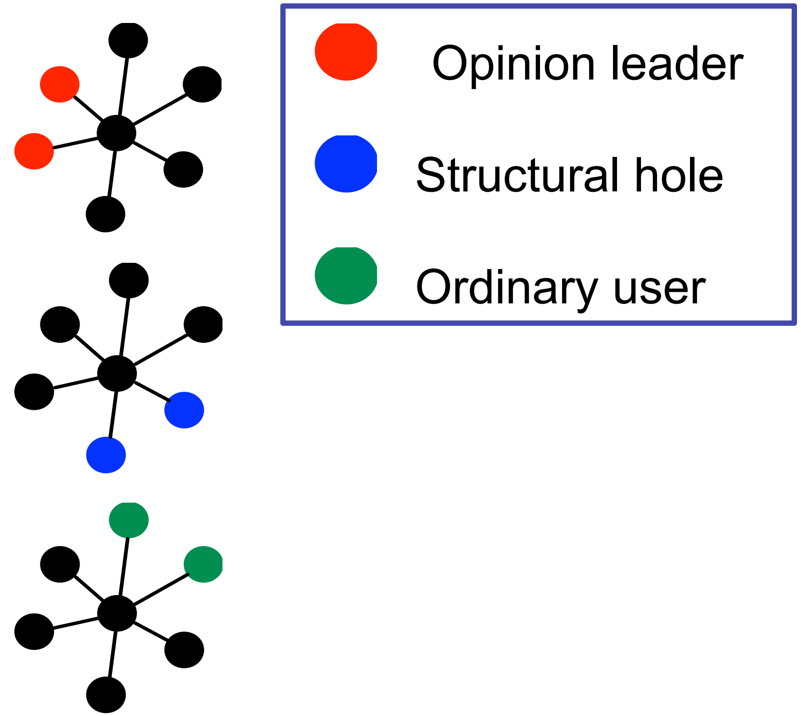
- Stage 1 - activation probability is 12 times higher than ordinary user
- Stage 2 - information overload[1]: 2-3 opinion leaders are sufficient to spread a piece of information throughout a community.
- Stage 3 - information everywhere: spreading the information becomes a social norm to adopt.

[1] Lazarsfeld, P. F.; Berelson, B.; and Gaudet, H. 1944. The peoples choice: How the voter makes up his mind in a presidential election. New York: Duell, Sloan and Pearce .



Influence Strength

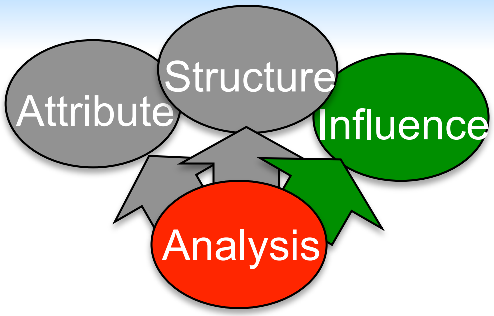




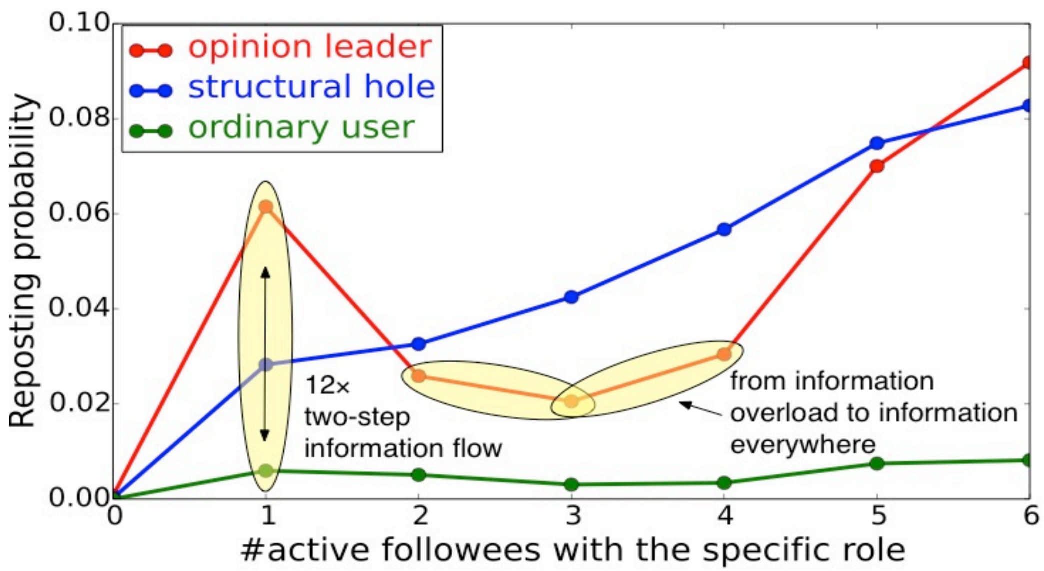
Opinion leader:

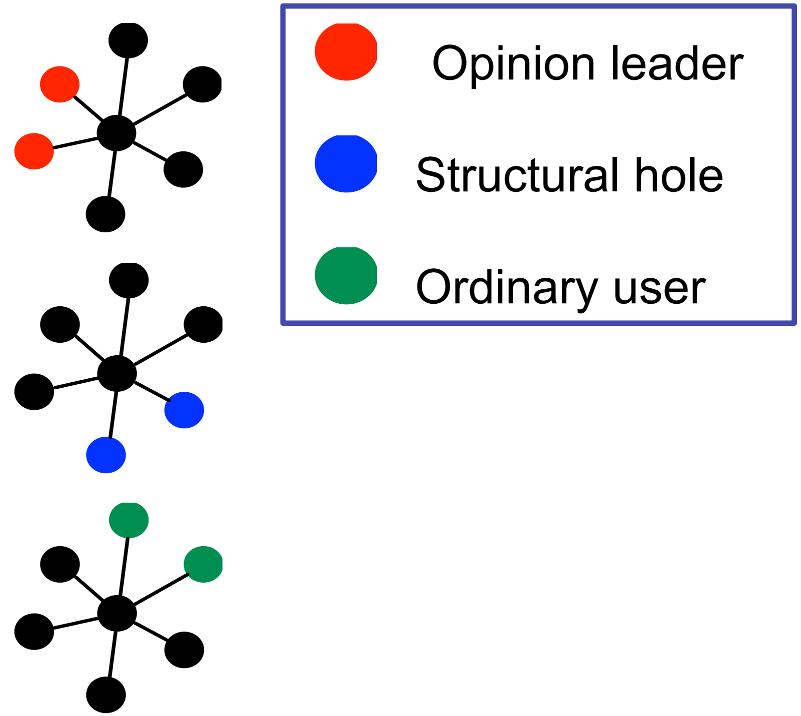
- Stage 1 - activation probability is 12 times higher than ordinary user
- Stage 2 - information overload[1]: 2-3 opinion leaders are sufficient to spread a piece of information throughout a community.
- Stage 3 - information everywhere: spreading the information becomes a social norm to adopt.

[1] Lazarsfeld, P. F.; Berelson, B.; and Gaudet, H. 1944. The peoples choice: How the voter makes up his mind in a presidential election. New York: Duell, Sloan and Pearce .



Influence Strength

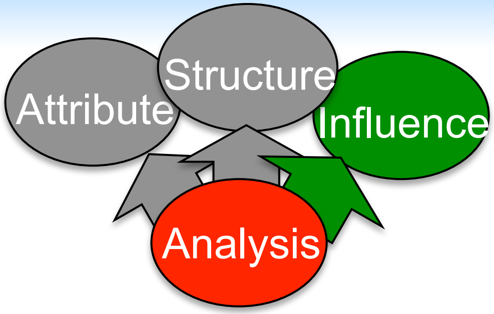




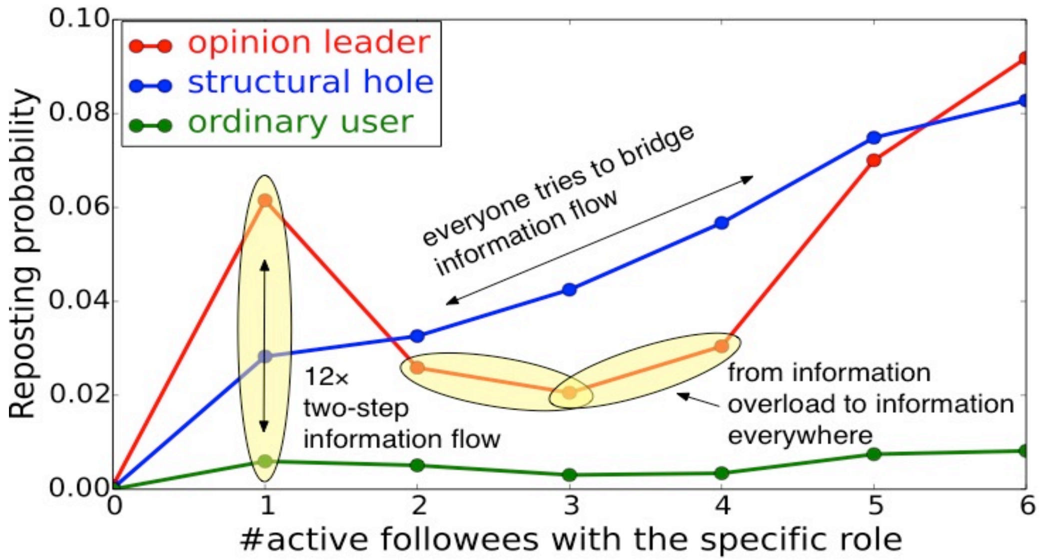
Opinion leader:

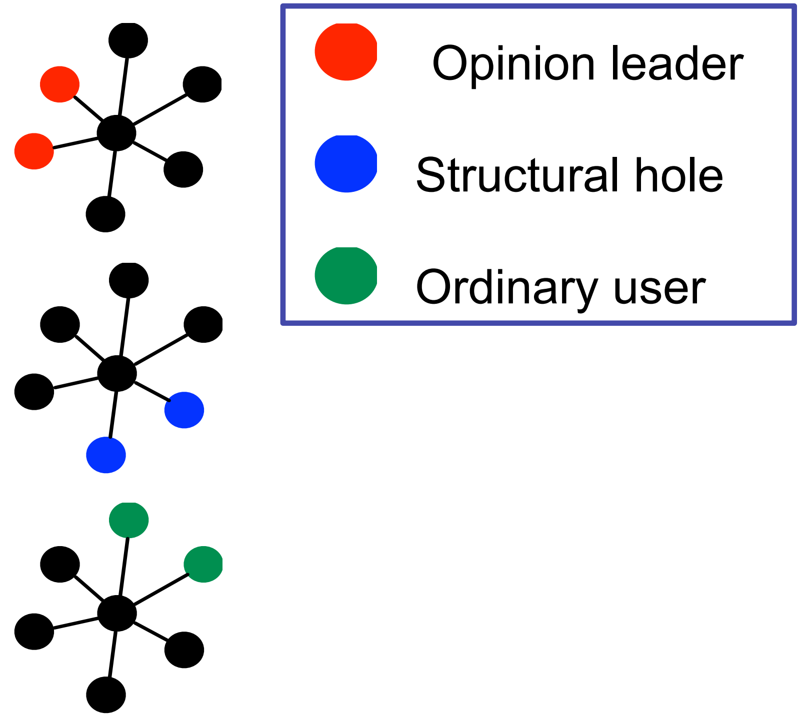
- Stage 1 - activation probability is 12 times higher than ordinary user
- Stage 2 - information overload[1]: 2-3 opinion leaders are sufficient to spread a piece of information throughout a community.
- Stage 3 - information everywhere: spreading the information becomes a social norm to adopt.

[1] Lazarsfeld, P. F.; Berelson, B.; and Gaudet, H. 1944. The peoples choice: How the voter makes up his mind in a presidential election. New York: Duell, Sloan and Pearce .



Influence Strength

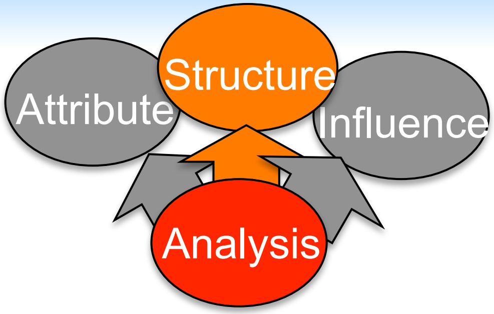




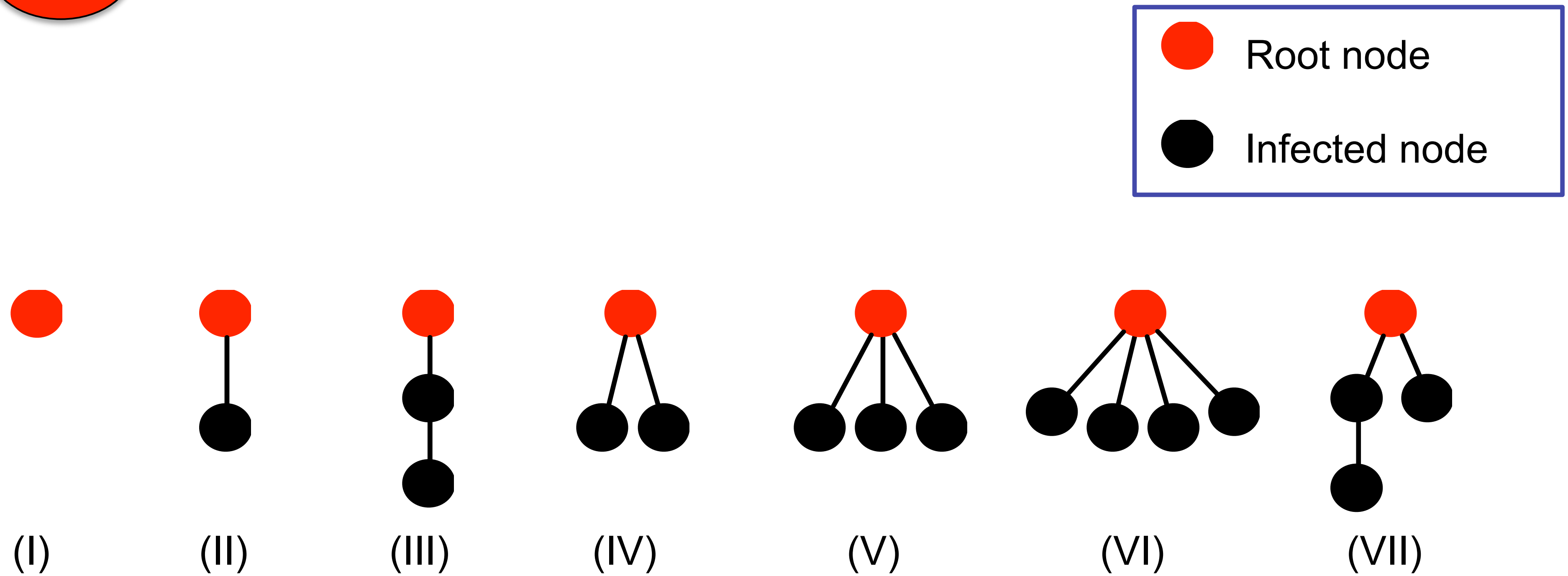
Structural hole spanners[2][3]:

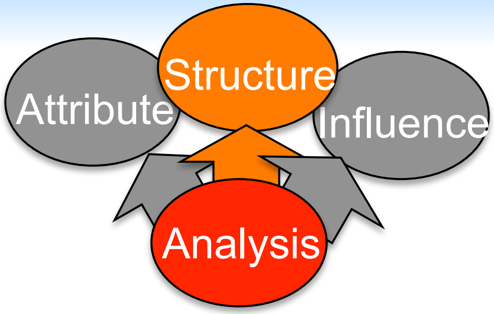
- SH tend to bring information that a certain community is rarely exposed to.
- Most users tries to bridge information flow between different groups.

[2] Burt, R. S. 2001. Structural holes versus network closure as social capital. Social capital: Theory and research 31–56.
[3] Burt, R. S. 2009. Structural holes: The social structure of competition . Harvard University Press.

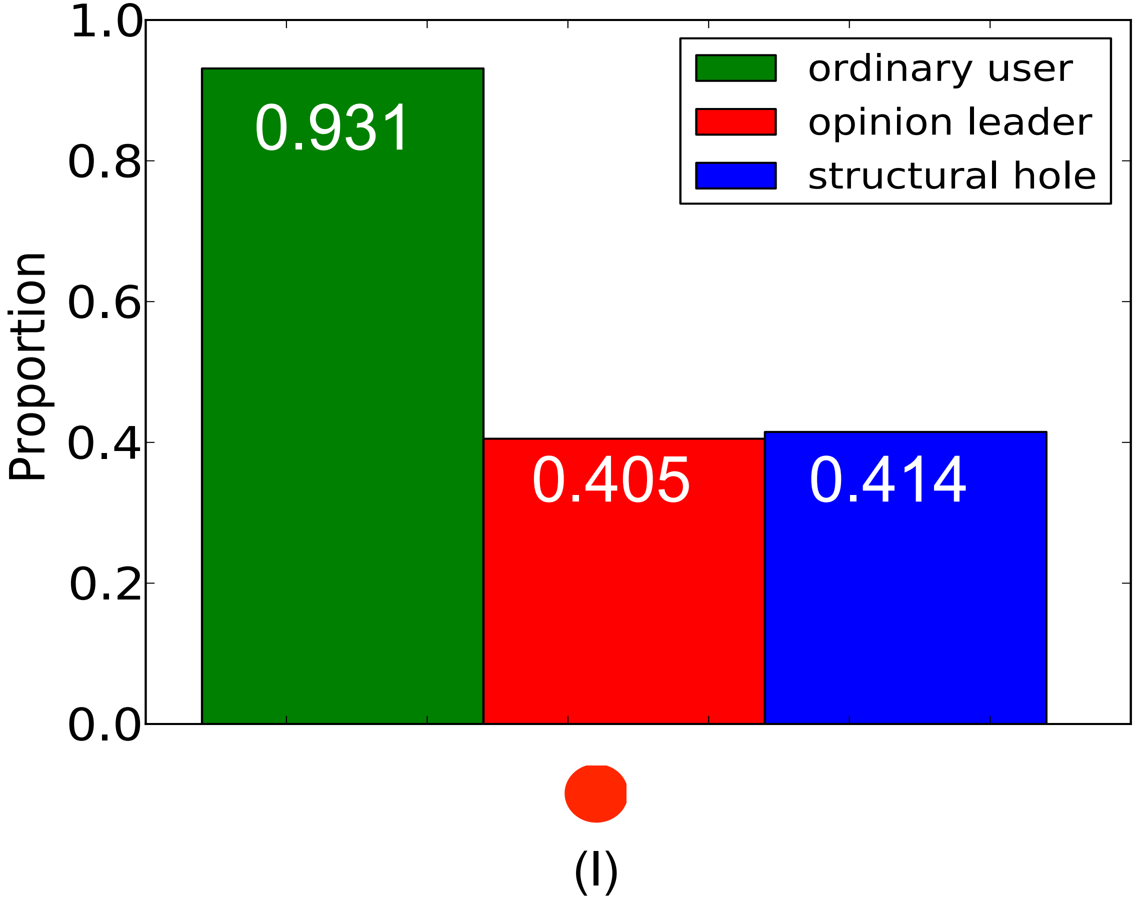


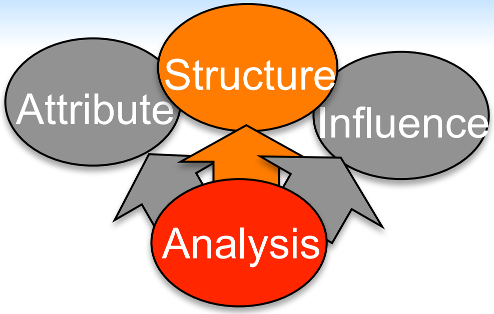
Atomic Diffusion Structure



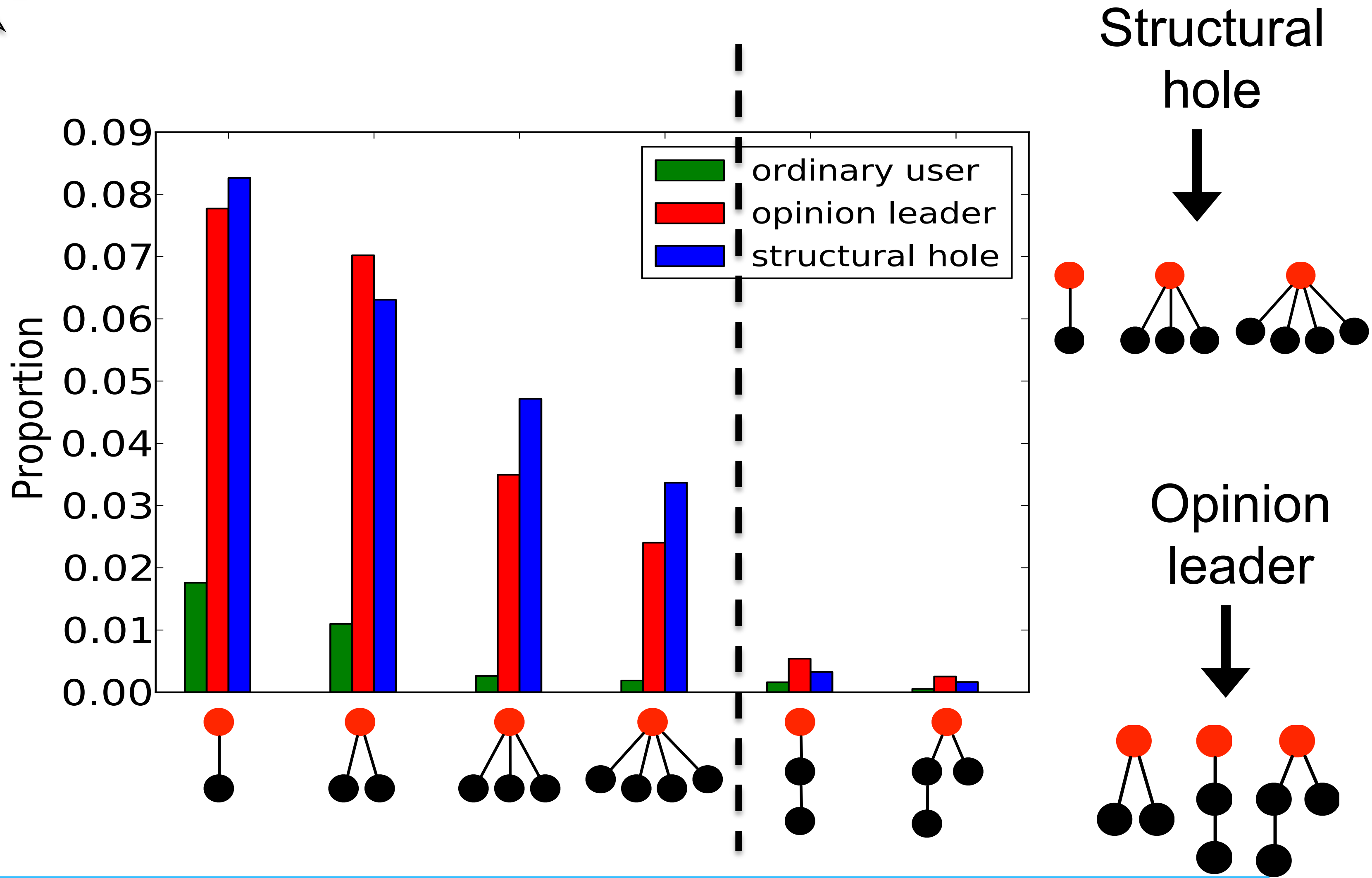


Atomic Diffusion Structure

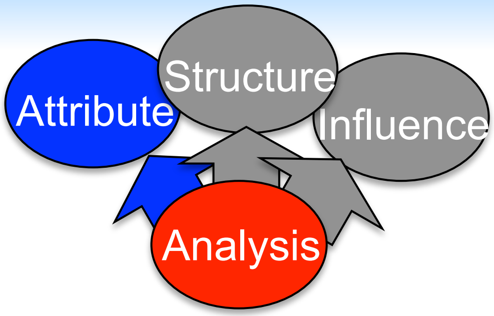




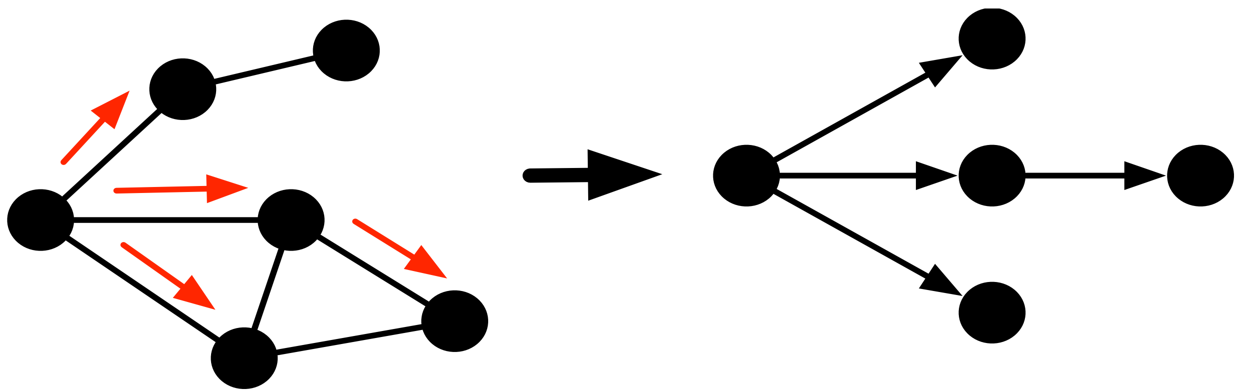
Atomic Diffusion Structure



Diffusion structures tend to be wide, and not too deep

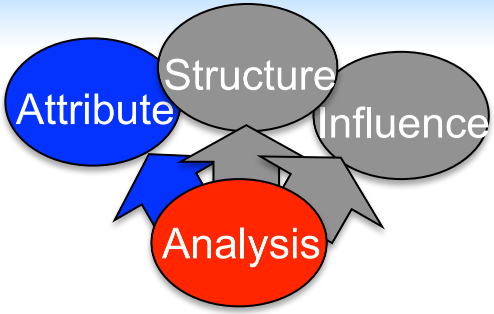


Formulation

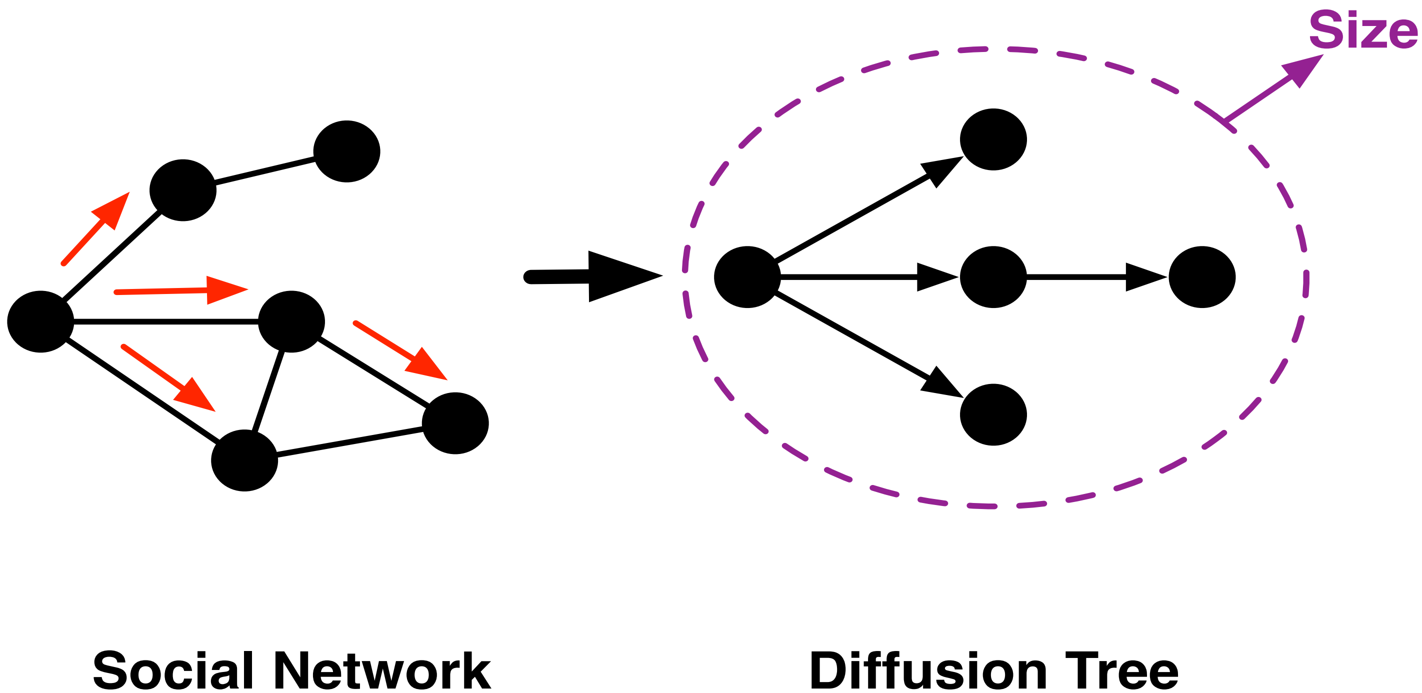


Social Network **Diffusion Tree**

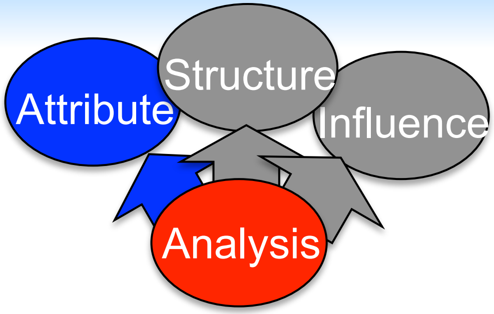
Definition 1. **Diffusion Tree.** In a given G, a diffusion tree of a message i comprises a set of 4-tuples: $\{(v', v, i, t)\}$, where each tuple (v', v, i, t) indicates that user v retweeted i from v' at time t. In a given tuple, $v' = -1$ *iff* v is the user who first posted i. In such case, the corresponding tuple is called the root of the diffusion tree.



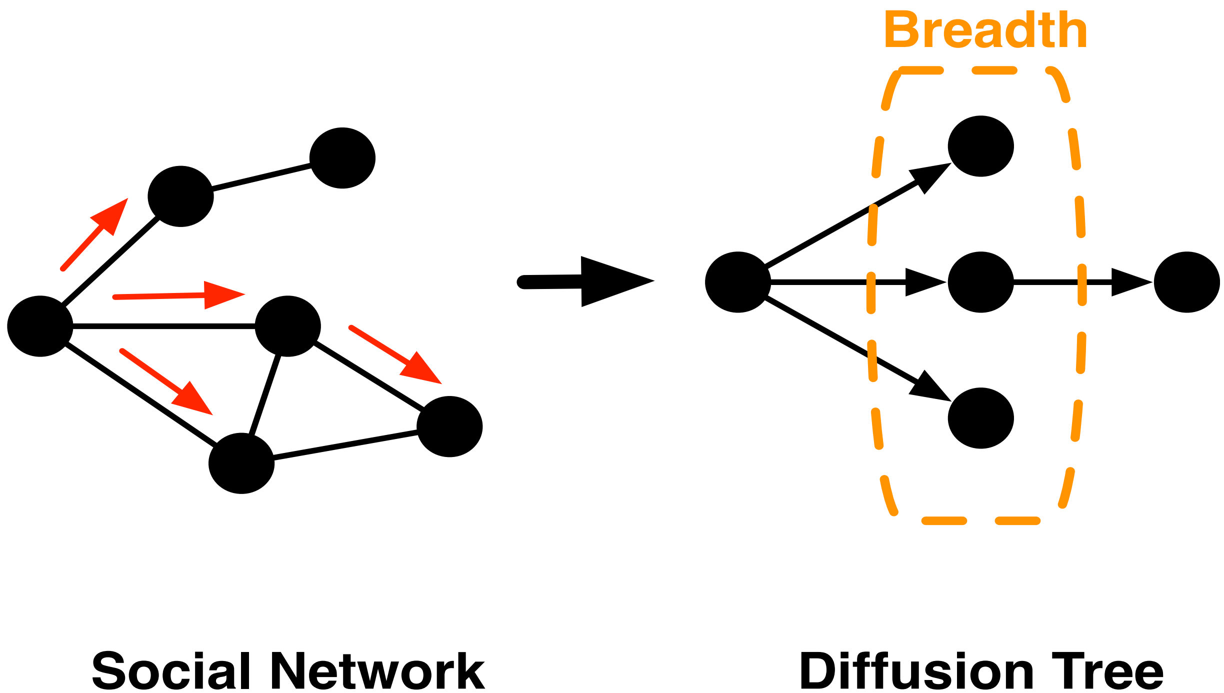
Formulation



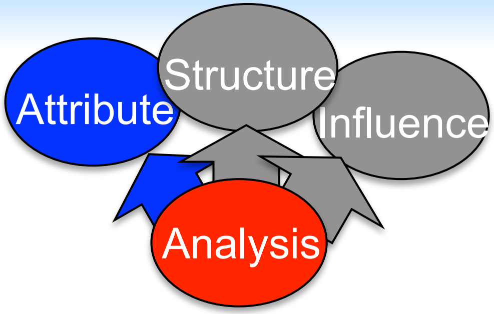
Diffusion size: how many users will receive the information



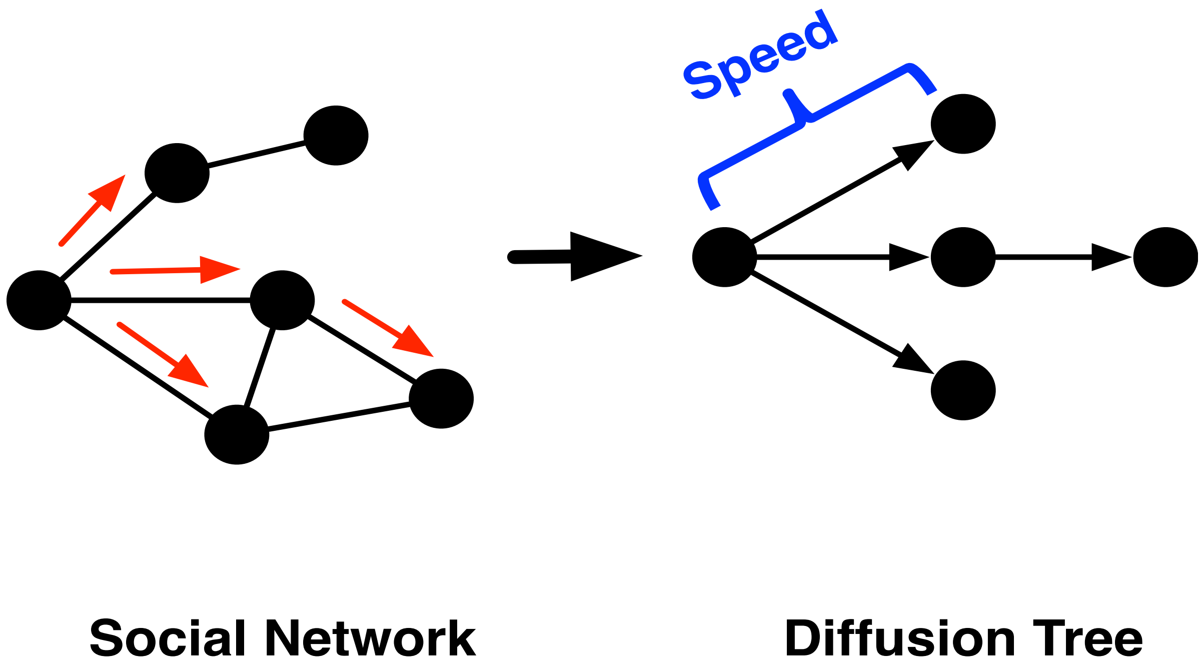
Formulation



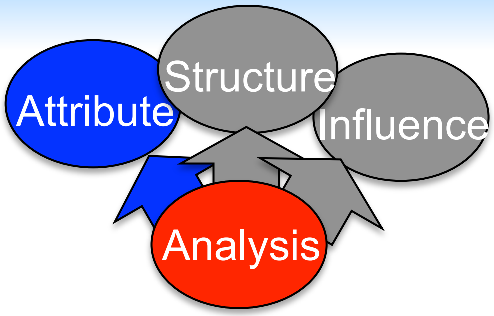
Diffusion breadth: how widely the information will propagate



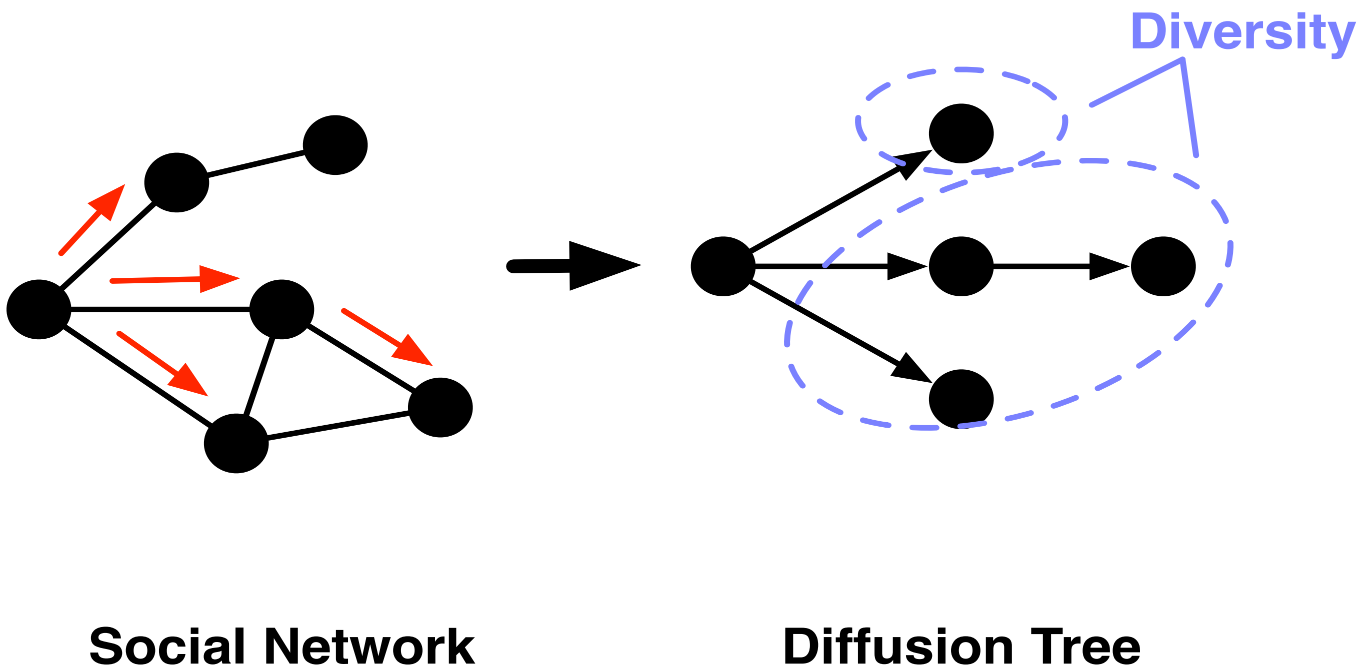
Formulation



Diffusion speed: how fast the information will propagate

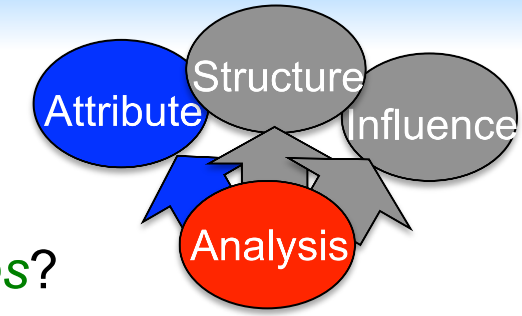


Formulation



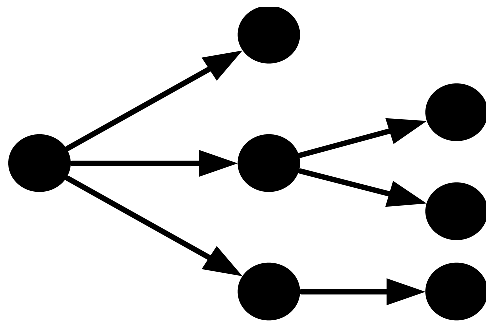
Diffusion diversity: how many communities will receive the information

Analysis Setup

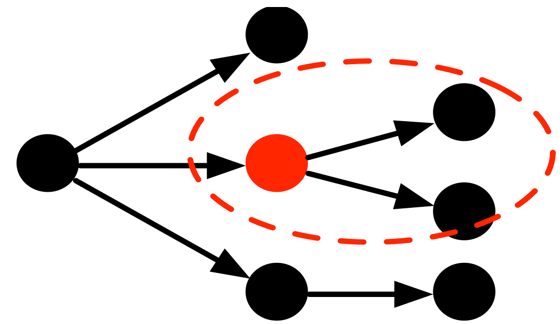


How different *social roles* influence different *diffusion attributes*?

Original diffusion tree

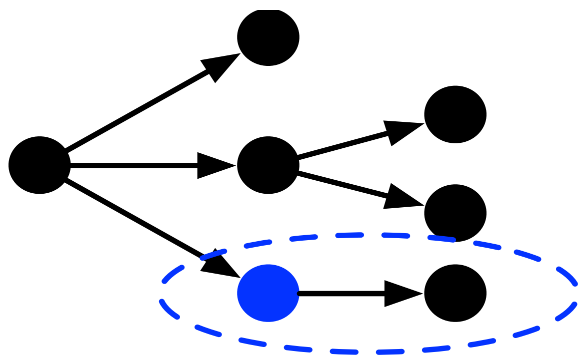


Opinion leader

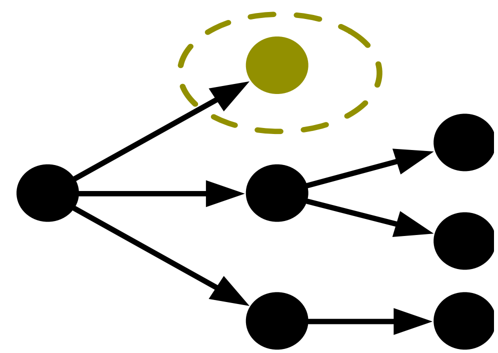


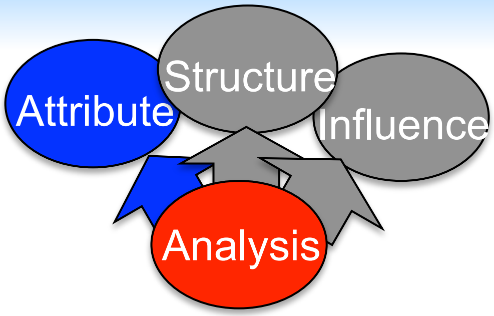
vs.

Structural hole spanner

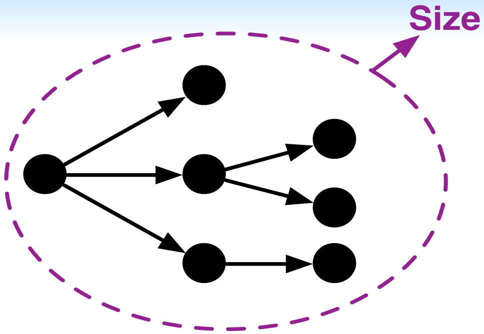


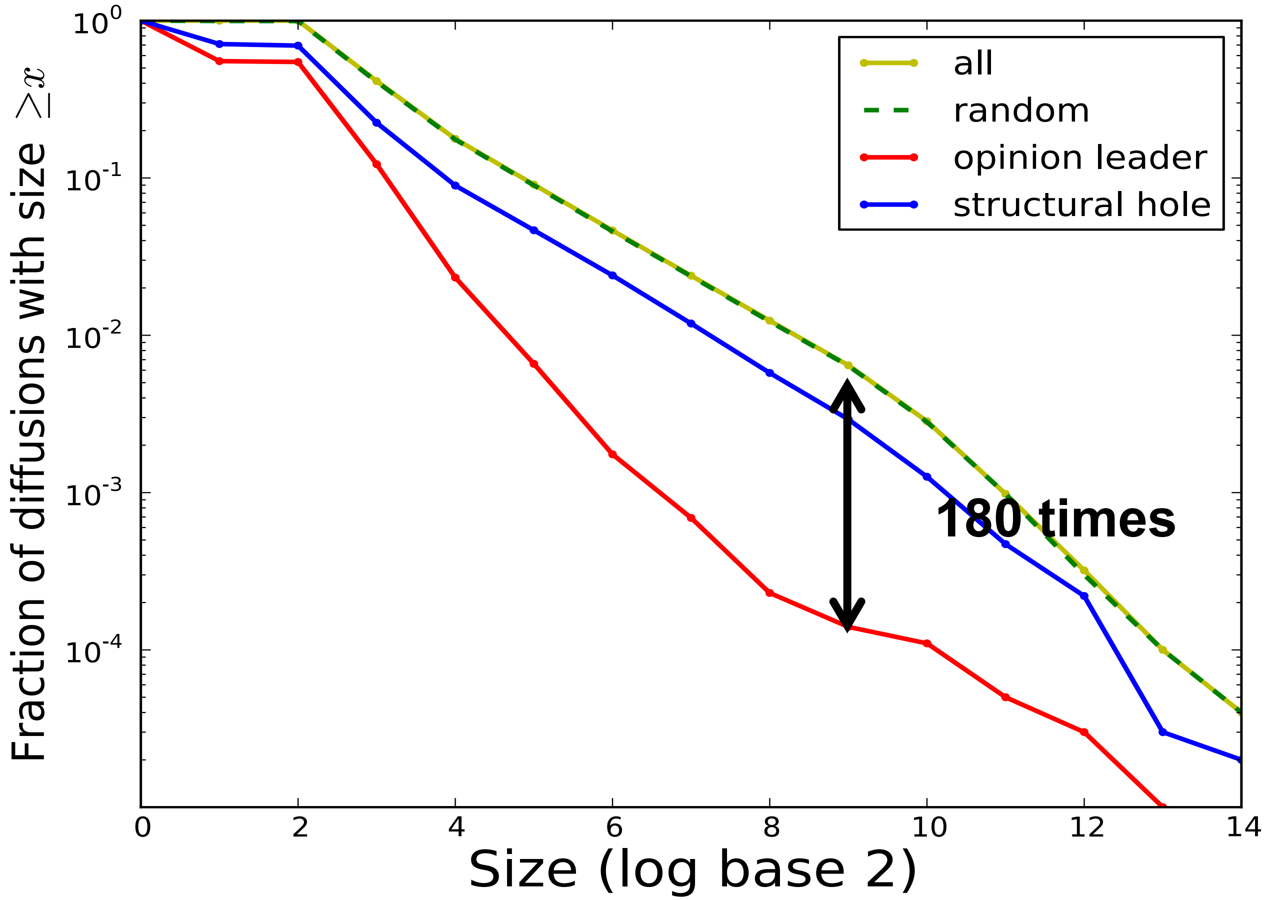
Random selected user



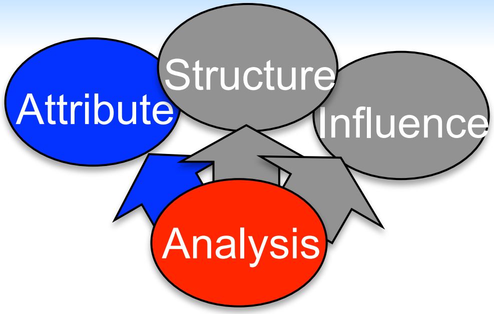


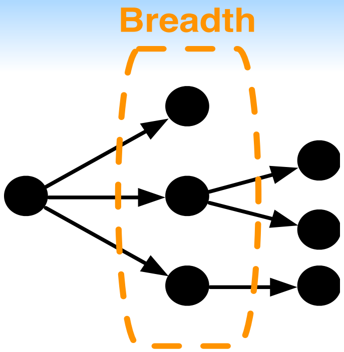
Diffusion Size

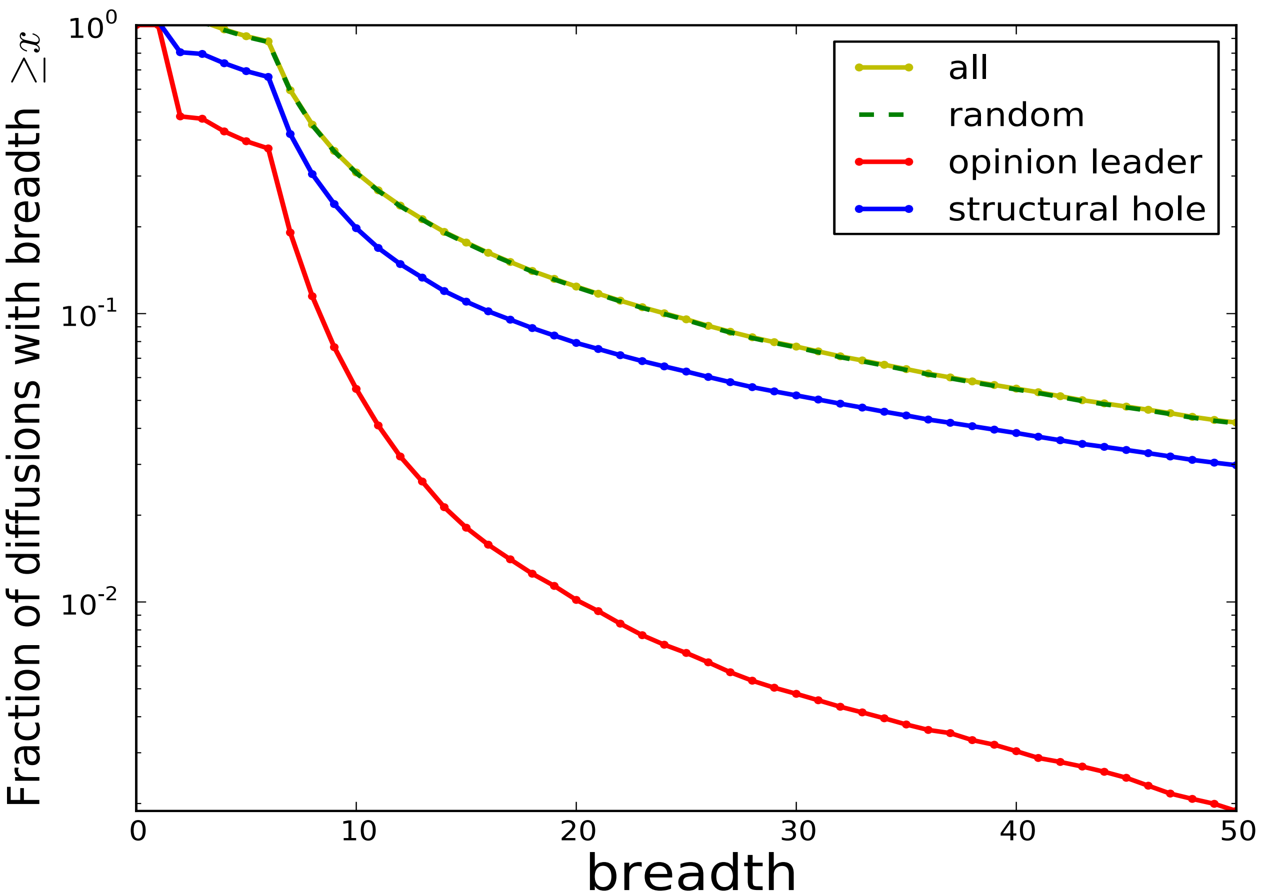


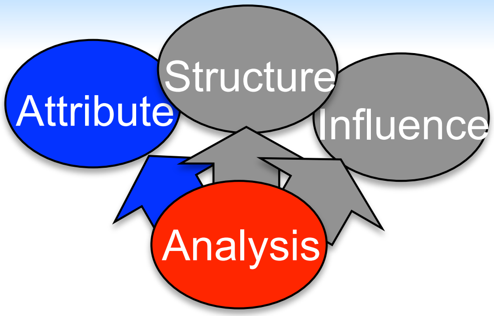


Diffusion Breadth

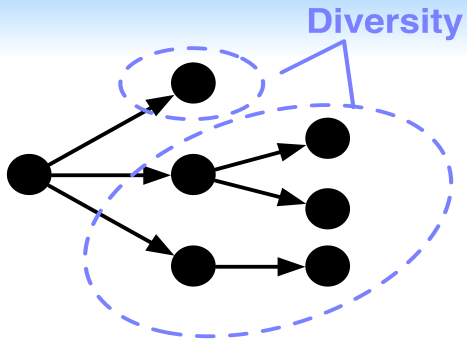


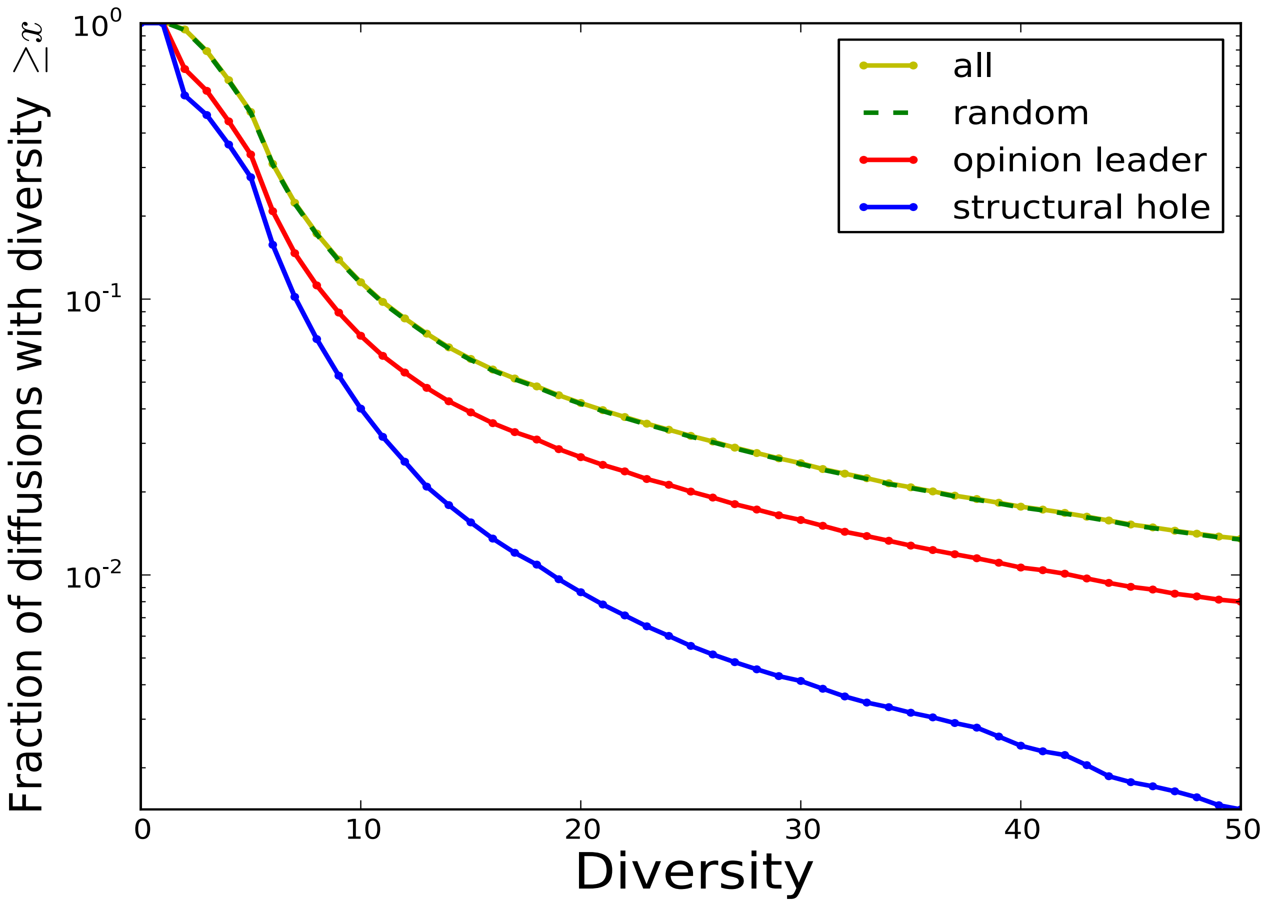


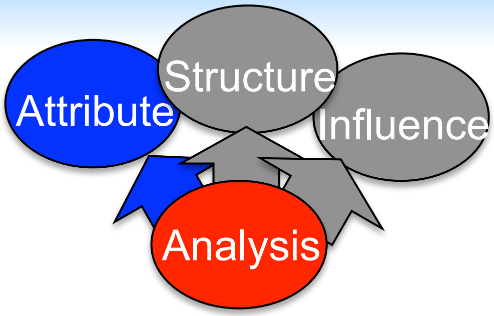




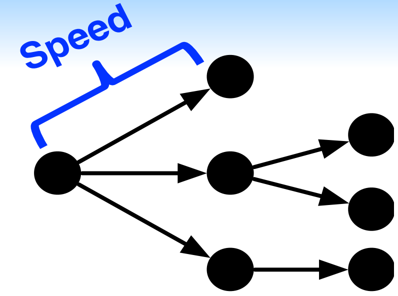
Diffusion Diversity

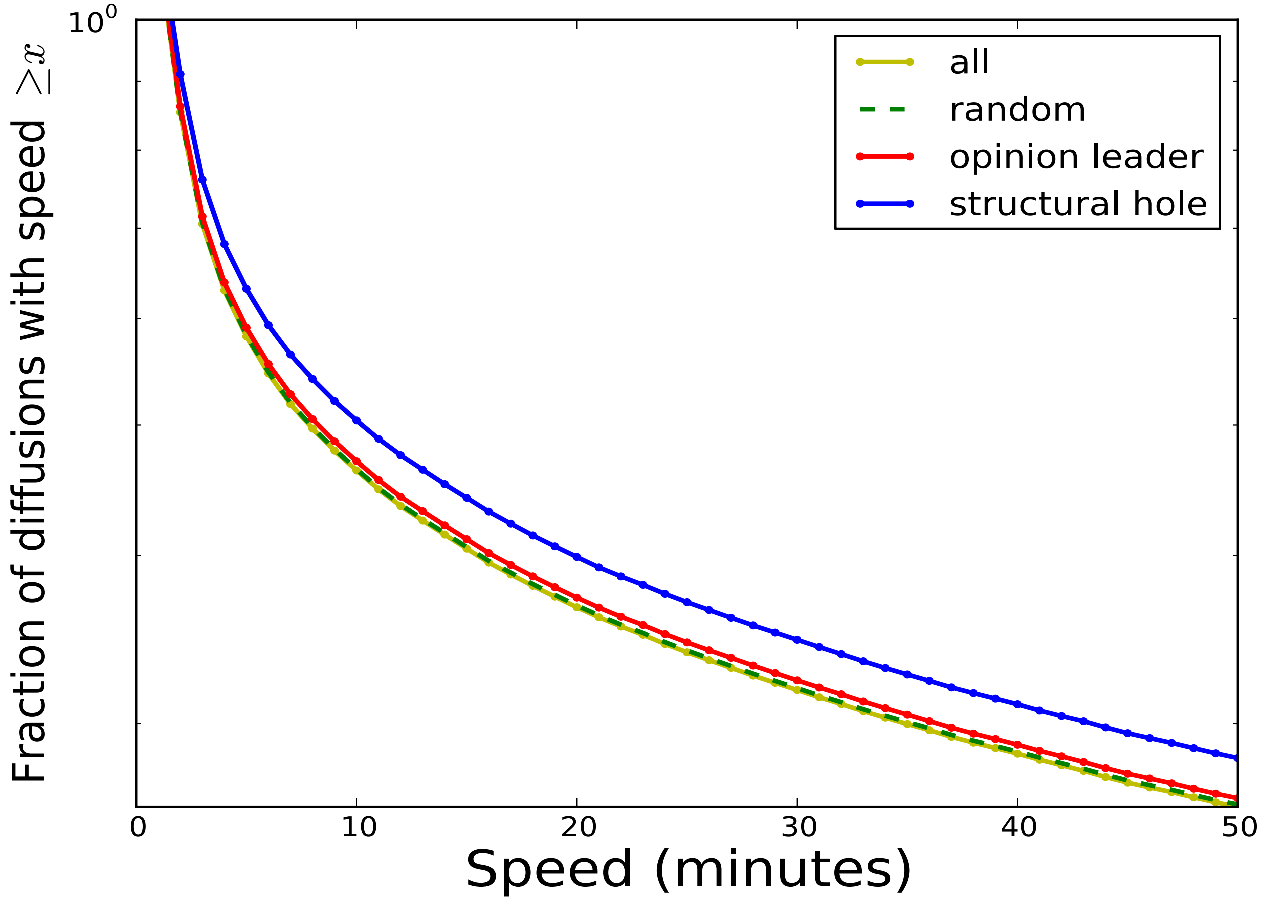


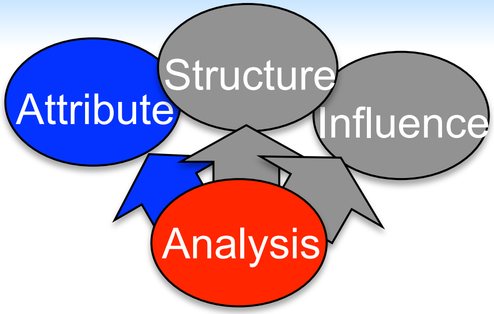




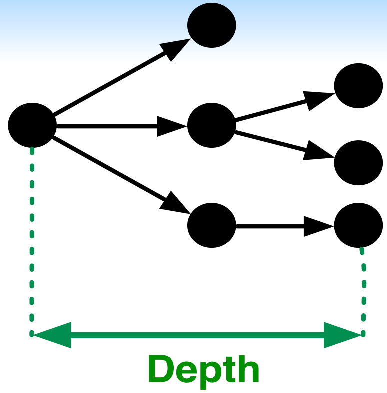
Diffusion Speed

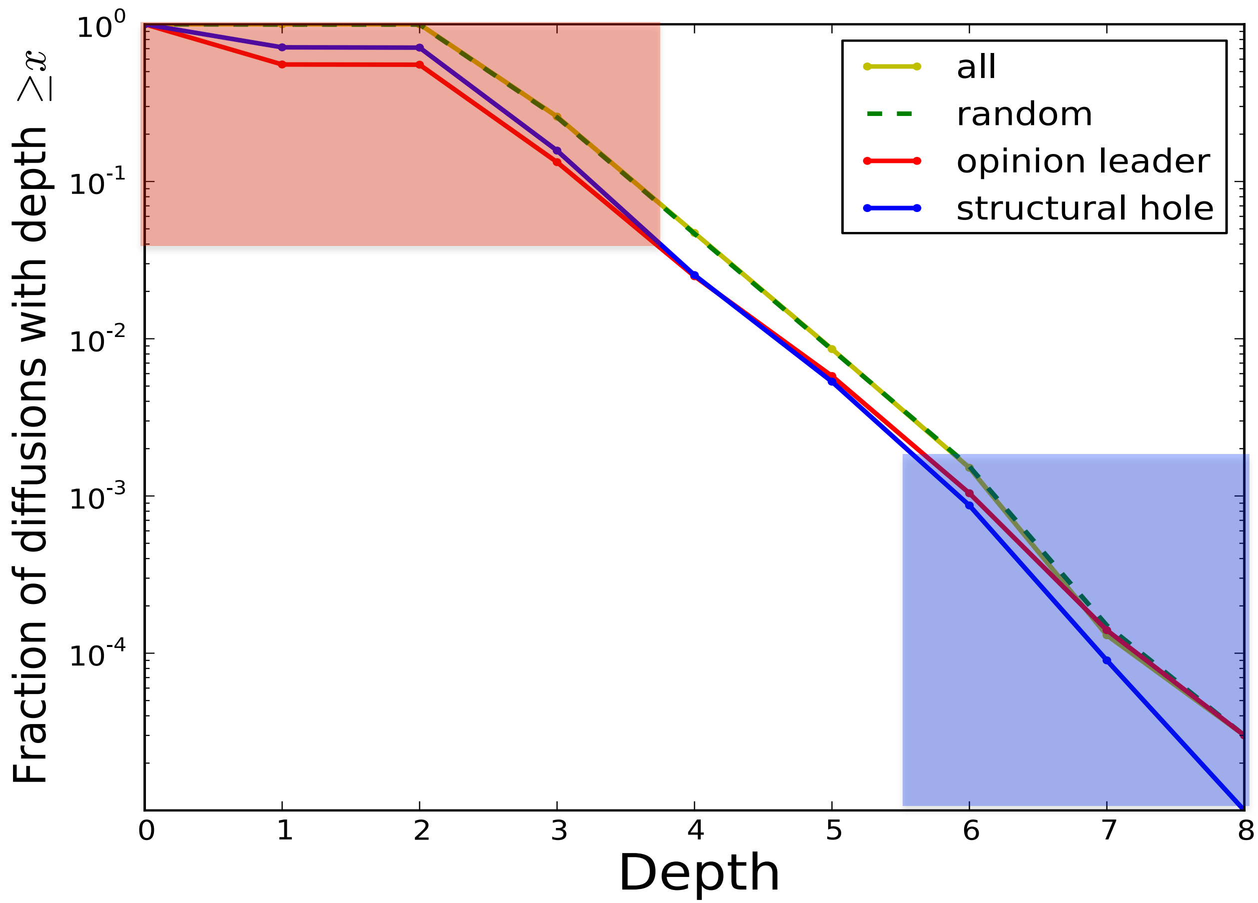


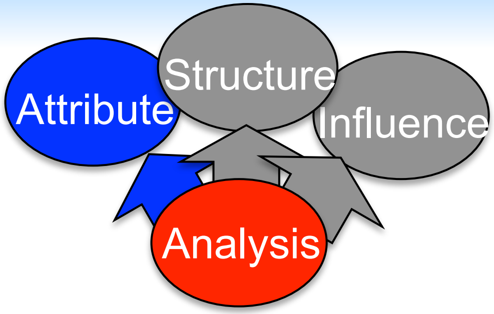




Diffusion Depth







Conclusion

- ***Opinion leaders*** are more influential on diffusion size & breadth;
- ***Structural hole spanners*** have more influence on diffusion diversity & speed;
- Diffusion depth is not sensitive to both opinion leaders and structural hole spanners.

How to better model information diffusion by leveraging social role information?

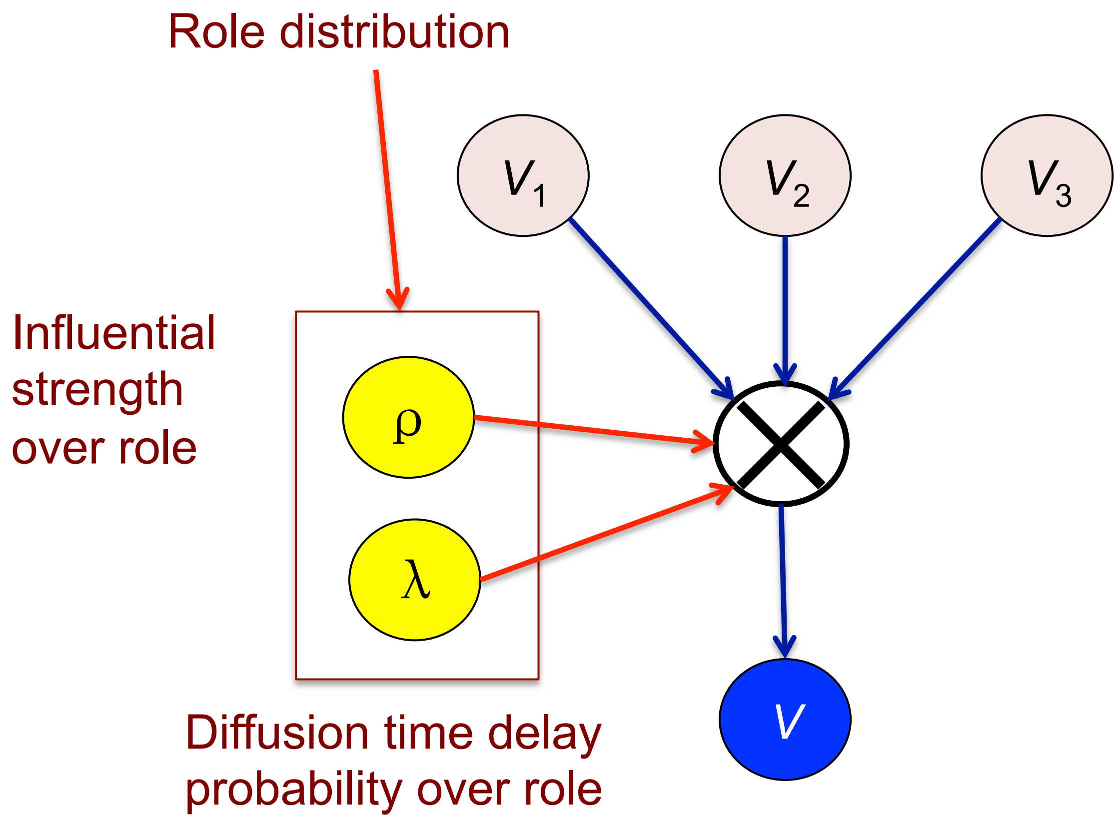
Given:

1. **A social network;**
2. **A set of historical diffusion trees.**

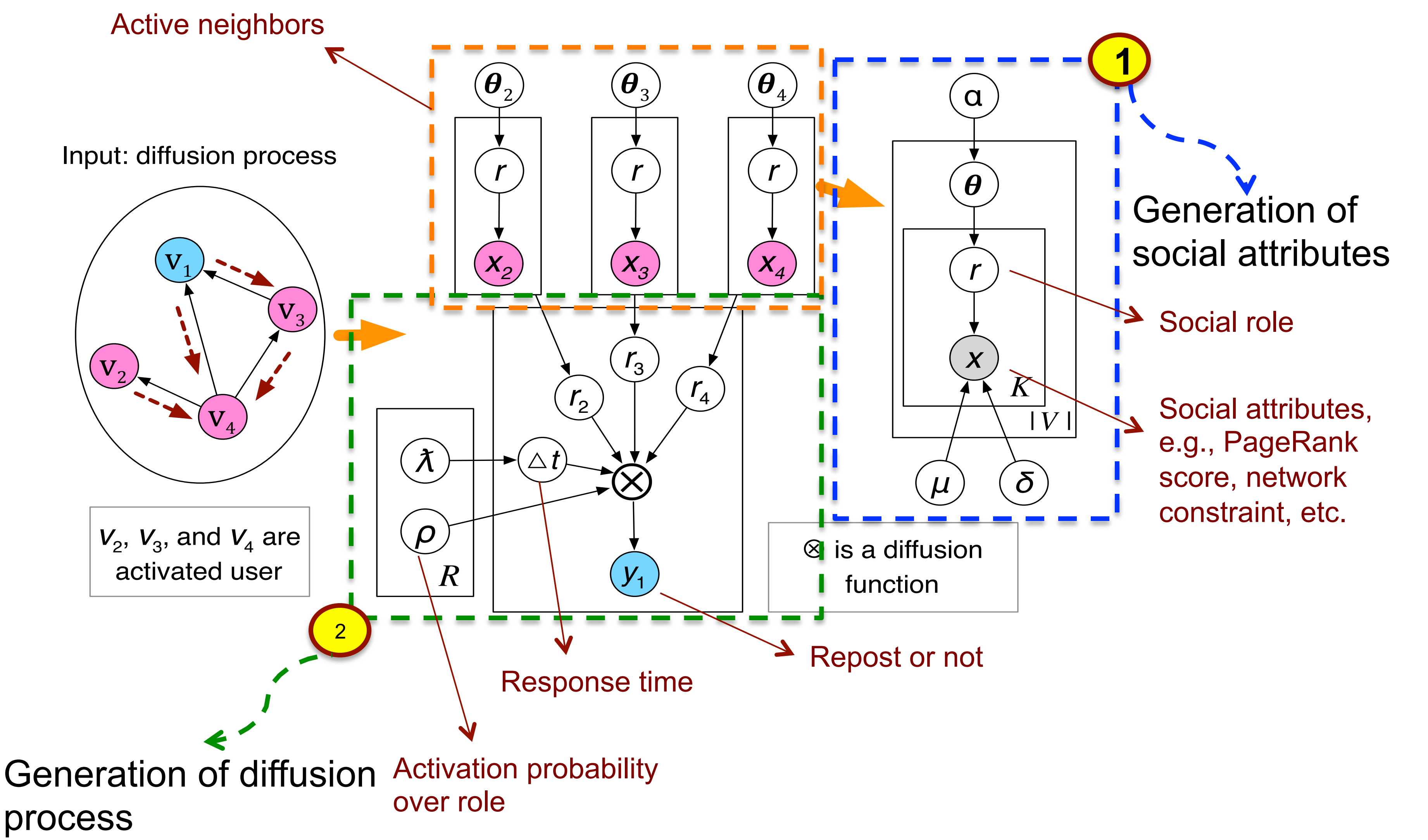
Goal:

1. **Model the diffusion process in future;**
2. **Infer social roles distributions of users.**

Model: General Idea



RAIN (Role Aware Information diffusioN)



RAIN: Objective Function

- Likelihood：

$$L = \prod_{i=1}^{I}\prod_{t=1}^{T}\prod_{v\in A_{it}} P(v\in A_{it}) \times \prod_{i=1}^{I}\prod_{v\notin D_{iT}} P(v\notin D_{iT})$$
$$\times \prod_{u\in V}\prod_{k=1}^{K} P(x_{uk}) \times \prod_{u\in V}\prod_{r=1}^{R} P(\theta_{ur}|\alpha)$$
$$\times \prod_{r=1}^{R}\{P(\rho_r|\beta)+P(\lambda_r|\gamma)\} \times \prod_{r=1}^{R}\prod_{k=1}^{K} P(\mu_{rk},\delta_{rk}|\tau)$$

The probability of user *v* adopting the information *i* at time *t*

$$P(v\in A_{it}) = \boxed{\sum_{\mathbf{z}^t_{i*v}} P(\mathbf{z}_{i*v})} - \prod_{u\in B(v)\cap D_{it-1}} \boxed{P(z^t_{iuv}=0)}$$
$$= \prod_{u\in B(v)\cap D_{it-1}} (\varphi^t_{iuv}+\varepsilon^t_{iuv}) - \prod_{u\in B(v)\cap D_{it-1}} \varepsilon^t_{iuv}.$$

All adoptions

Failed adoptions

The probability of user *v* never adopts the information *i*

$$P(v\notin D_{iT}) = \prod_{u\in B(v)\cap D_{iT}} \sum_r (1-\rho_r)\theta_{ur}.$$

Assumption here:
T >> the last observed timestamp

The probability of user *v* with the social attributes x_{vk}

$$P(x_{uk}) = \sum_r \sqrt{\frac{\delta_{rk}}{2\pi}} \exp\{-\frac{\delta_{rk}(x_{uk}-\mu_{rk})^2}{2}\}\theta_{ur}.$$

A mixture of Gaussian

Priors to model parameters

Model Learning

Gibbs Sampling:

- Sample latent role *r* for user u's each social attribute

$$P(r_{uk}|\mathbf{r}_{\neg uk},\mathbf{x}) = \frac{P(\mathbf{x},\mathbf{r})}{P(\mathbf{x}_{\neg uk},\mathbf{r}_{\neg uk})} = \frac{n_{ur_{uk}}^{\neg uk}+\alpha}{\sum_r (n_{ur}^{\neg uk}+\alpha)} \frac{\Gamma(\tau_2 + \frac{n_{r_{uk}k}}{2})}{\Gamma(\tau_2 + \frac{n_{r_{uk}k}^{\neg uk}}{2})} \times \frac{\sqrt{(\tau_1 + n_{r_{uk}k}^{\neg uk})}\eta(n_{r_{uk}k}^{\neg uk}, \bar{x}_{r_{uk}k}^{\neg uk}, s_{r_{uk}}^{\neg uk})}{\sqrt{(\tau_1 + n_{r_{uk}k})}\eta(n_{r_{uk}k}, \bar{x}_{r_{uk}k}, s_{r_{uk}})},$$

- Sample role *r*, time delay *t*, and activation result *z* for each adoption

$$\begin{aligned}
&P(r_{iuv},\Delta t_{iuv}, z_{iuv}|\mathbf{r}_{\neg iuv},\Delta\mathbf{t}_{\neg iuv},\mathbf{z}_{\neg iuv},\mathbf{y}) \\
&= \frac{P(\mathbf{r},\Delta\mathbf{t},\mathbf{z},\mathbf{y})}{P(\mathbf{r}_{\neg iuv},\Delta\mathbf{t}_{\neg iuv},\mathbf{z}_{\neg iuv},\mathbf{y}_{\neg iuv})} \\
&= \frac{n_{ur_{iuv}}^{\neg iuv}+\alpha}{\sum_r (n_{ur}^{\neg iuv}+\alpha)} \times \frac{n_{z_{iuv}r_{iuv}}^{\neg iuv} + \beta_1^{z_{iuv}}\beta_0^{1-z_{iuv}}}{n_{1r_{iuv}}^{\neg iuv}+\beta_1+n_{0r_{iuv}}^{\neg iuv}+\beta_0} \\
&\times \frac{(n_{r_{iuv}}^{\neg iuv}+\gamma_1)\prod_{t=0}^{\Delta t-2}(s_{r_{iuv}}^{\neg iuv}-n_{r_{iuv}}^{\neg iuv}+\gamma_0+t)}{\prod_{t=0}^{\Delta t-1}(\gamma_1+s_{r_{iuv}}^{\neg iuv}+\gamma_0+t)} \times \Phi,
\end{aligned}$$

- Update model parameters according to sampling results

Input: the hyper-parameters α, β, γ, and τ, the number of social roles R, a social network G along with each user's social attribute $\mathbf{x}_v$, and a set of diffusion trees.

```
foreach user u ∈ V do
    Initialize θ_u randomly;
end
for r = 1 to R do
    Initialize ρ_r and λ_r randomly;
end
repeat
    % sampling process;
    foreach user u ∈ V do
        for k = 1 to K do
            Draw a latent variable r, which is associated
            with x_uk, according to P(r_uk|r_¬uk, x) (Eq. 7);
        end
    end
    foreach 4-tuple (u, v, i, t) in each diffusion tree do
        Draw latent variables (t, r, z) according to
        P(r_iuv, Δt_iuv, z_iuv|r_¬iuv, Δt_¬iuv, z_¬iuv, y) (Eq. 9);
    end
    % parameter update;
    for r = 1 to R do
        Update λ_r and ρ_r according to Eq. 10;
        foreach user u ∈ V do
            Update θ_ur according to Eq. 10;
        end
        for k = 1 to K do
            Update μ_rk and δ_rk according to Eq. 11
        end
    end
until Convergence;
```

Retweet Prediction

Table 2: Performance of repost prediction on several topics.

Topic	Method	P@10	P@50	P@100	MAP
Campus	C				
	S				
	IC				
	R				
Horoscope	C				
	S				
	IC				
	R				
Movie	C				
	S				
	IC				
	R				
History	C				
	S				
	IC				
	R				
Society	C				
	S				
	IC				
	R				
Health	C				
	S				
	IC				
	R				
Political	C				
	S				
	IC Model	[illegible]	[illegible]	[illegible]	[illegible]
	RAIN	**0.216**	**0.164**	**0.130**	**0.239**
Travel	Count	0.142	0.056	0.031	0.103
	SVM	0.094	0.048	0.032	0.128
	IC Model	**0.206**	0.120	0.098	0.254
	RAIN	0.194	**0.159**	**0.126**	**0.260**

Goal: predict whether a user will repost a particular post

Data: a complete Tencent Weibo data on Nov. 1-3, 2012

- Posts are categorized based on topics: ***campus, constellation, movie, history, society, health, political, and travel***
- Posts on Nov.1-2 as train data, Nov. 3 as test data

Retweet Prediction

Table 2: Performance of repost prediction on several topics.

Topic	Method	P@10	P@50	P@100	MAP
Campus	Count	0.028	0.010	0.006	0.068
	SVM				
	IC Model				
	RAIN				
Horoscope	Count				
	SVM				
	IC Model				
	RAIN				
Movie	Count				
	SVM				
	IC Model				
	RAIN				
History	Count				
	SVM				
	IC Model				
	RAIN				
Society	Count				
	SVM				
	IC Model				
	RAIN				
Health	Count				
	SVM				
	IC Model				
	RAIN				
Political	Count				
	SVM				
	IC Model				
	RAIN				
Travel	Count				
	SVM				
	IC Model				
	RAIN				

Baselines:

Count: ranks users by the number of active followees

SVM: Support Vector Machine, majorly considers features as

- *#active followers*
- *#active followees*
- *#whether the user have reposted similar messages*

IC Model: traditional IC model with fitted parameters[1]

RAIN: **R**ole **A**ware **IN**formation diffusion

Evaluation Metrics:

Precision@K (K=10, 50, 100)

Mean Average Precision (MAP)

[1] Kimura, M.; Saito, K.; Ohara, K.; and Motoda, H. 2011. Learning information diffusion model in a social network for predicting influence of nodes. Intelligent Data Analysis 15(4):633–652.

Retweet Prediction

Table 2: Performance of repost prediction on several topics.

Topic	Method	P@10	P@50	P@100	MAP
Campus	Count	0.028	0.010	0.006	0.068
	SVM	0.098	0.045	0.032	0.127
	IC Model	**0.231**	0.142	0.102	0.259
	RAIN	0.228	**0.145**	**0.106**	**0.263**
Horoscope	Count	0.019	0.010	0.006	0.005
	SVM	0.124	**0.162**	0.088	**0.263**
	IC Model	0.149	0.111	0.098	0.125
	RAIN	**0.171**	0.121	**0.102**	0.130
Movie	Count	0.015	0.007	0.004	0.009
	SVM	0.094	0.111	0.060	0.199
	IC Model	0.227	0.147	**0.147**	0.236
	RAIN	**0.229**	**0.173**	0.144	**0.238**
History	Count	0.191	0.056	0.033	0.096
	SVM	0.154	0.051	0.030	0.221
	IC Model	0.206	0.134	**0.135**	0.230
	RAIN	**0.225**	**0.171**	0.134	**0.262**
Society	Count	0.245	0.058	0.029	0.156
	SVM	0.100	0.023	0.012	0.122
	IC Model	0.171	0.131	**0.109**	0.198
	RAIN	**0.176**	**0.140**	0.106	**0.204**
Health	Count	0.041	0.008	0.005	0.035
	SVM	0.164	0.064	0.039	**0.197**
	IC Model	0.169	0.113	0.096	0.162
	RAIN	**0.175**	**0.134**	**0.115**	0.185
Political	Count	0.019	0.005	0.003	0.007
	SVM	0.104	0.077	0.039	0.176
	IC Model	0.209	0.132	0.102	0.224
	RAIN	**0.216**	**0.164**	**0.130**	**0.239**
Travel	Count	0.142	0.056	0.031	0.103
	SVM	0.094	0.048	0.032	0.128
	IC Model	**0.206**	0.120	0.098	0.254
	RAIN	0.194	**0.159**	**0.126**	**0.260**

Comparison Results:

- **Count: performs worst due to the lack of supervised information.**
- **SVM: performs well on *local topics* but falls short on *global topics*.**
- **IC Model: suffers from *model complexity*.**
- **RAIN: improves the performance *+32.6%* in terms of MAP by reducing model complexity.**

Diffusion Scale Prediction

- We predict the ***scale*** of a diffusion process
  - X-axis: the number of reposts
  - Y-axis: the proportion of original posts with particular number of reposts

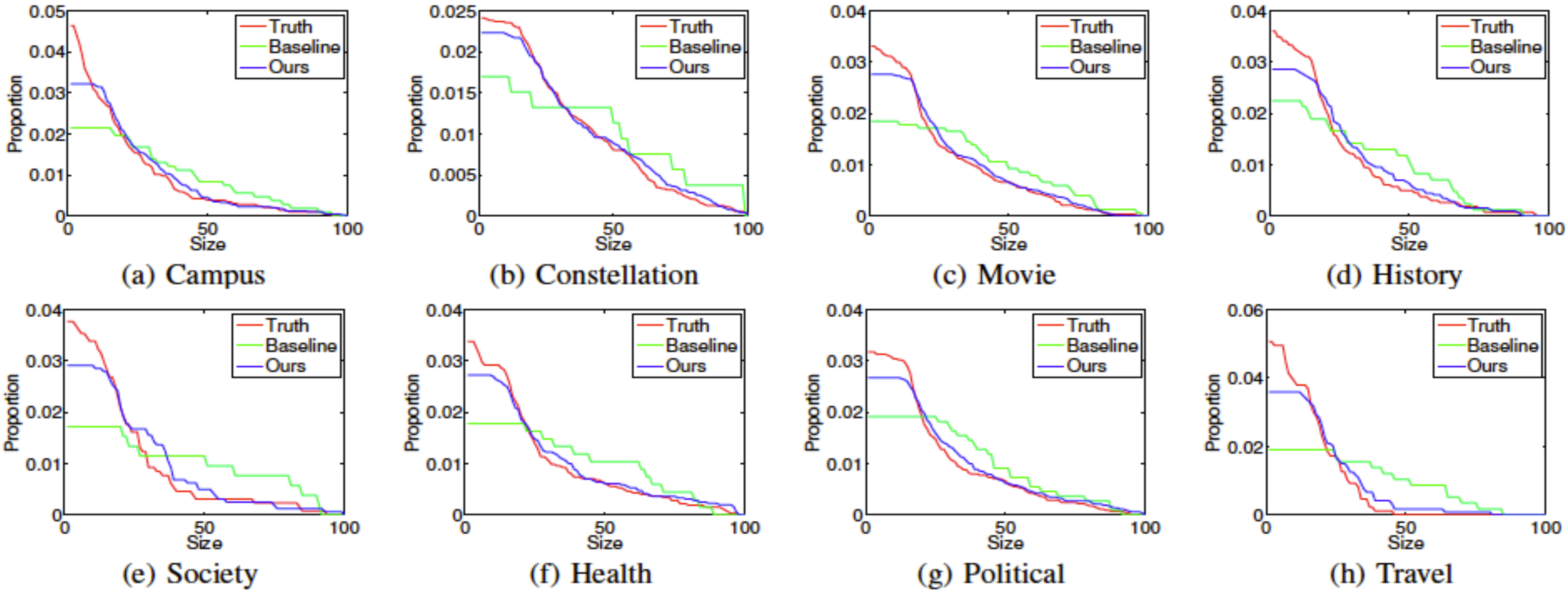


Figure 8: Diffusion scale distributions of the different topics in the test set.

Diffusion Duration Prediction

- We predict the ***duration*** of a diffusion process
  - X-axis: the time interval between the first and last posts
  - Y-axis: the proportion of original posts with particular time interval

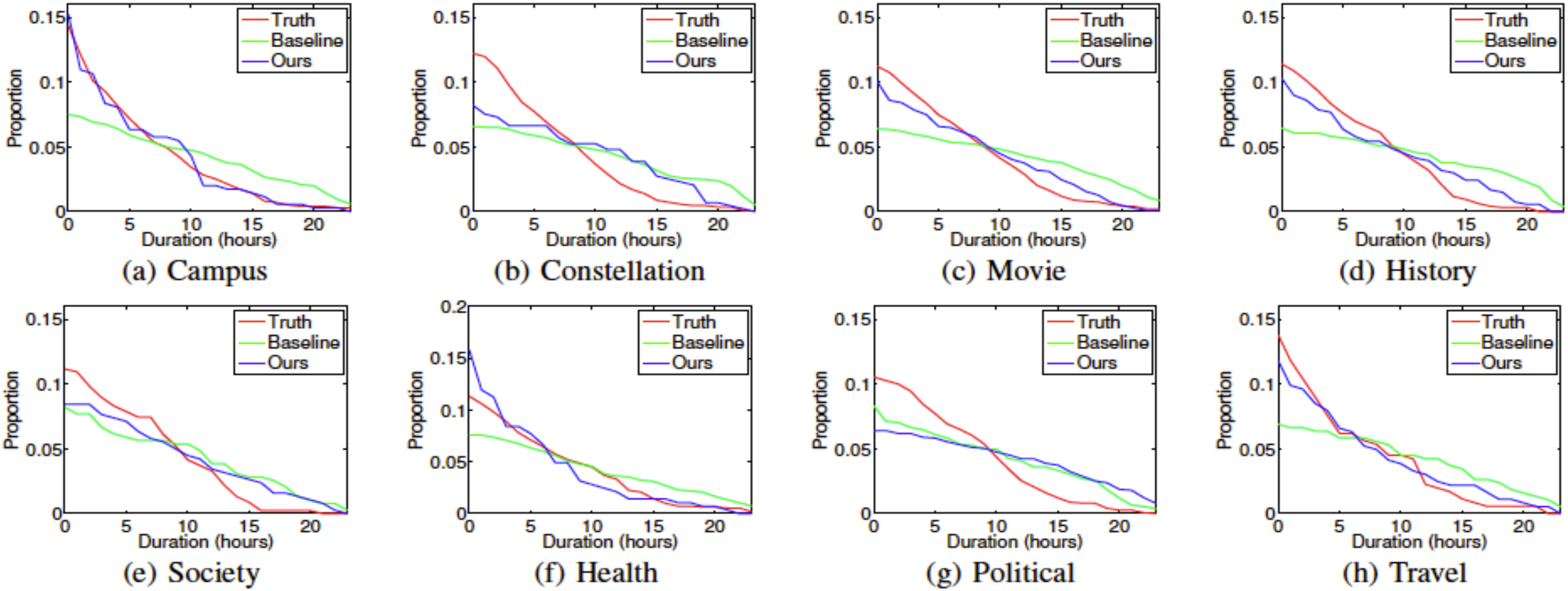


Figure 9: Diffusion duration distributions of the different topics in the test set.

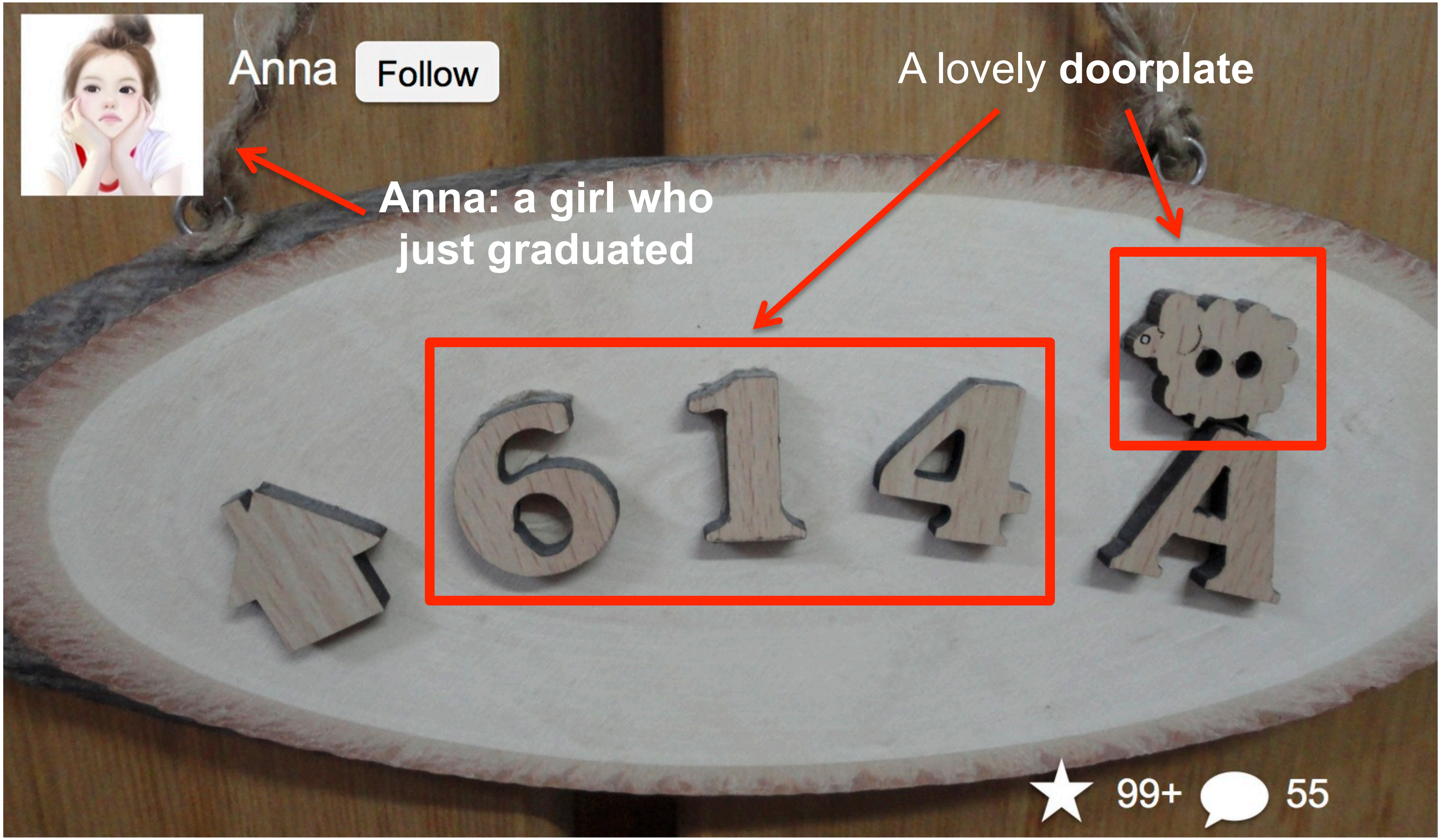
Part II:
User Emotion Influence and Influence based Network Embedding

How Do User Emotions Diffuse in Social Networks?

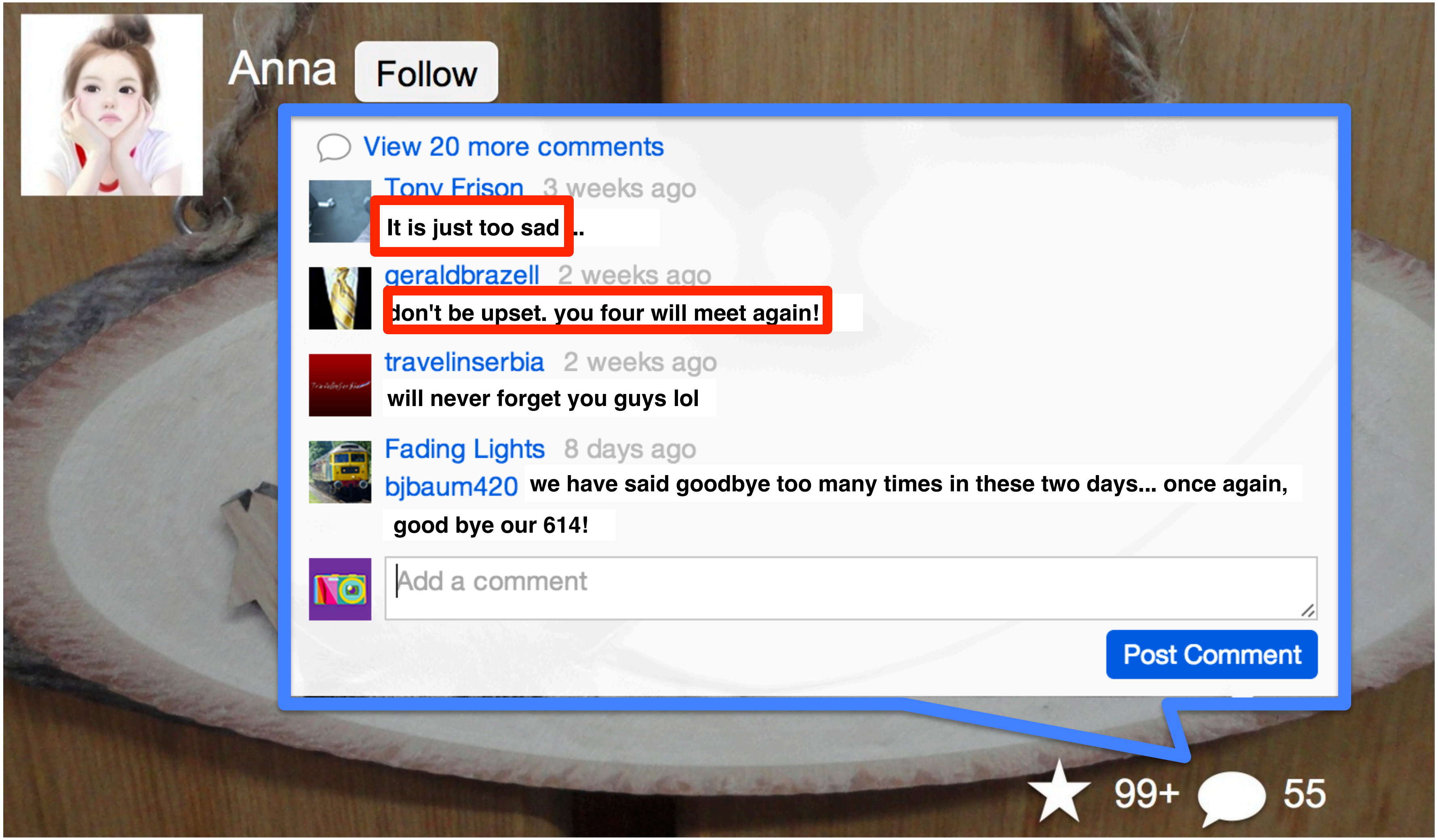
Yang Yang, Jia Jia, Boya Wu, and Jie Tang. **Social Role-Aware Emotion Contagion in Image Social Networks.** AAAI, 2016.

Yang Yang, Jia Jia, Shumei Zhang, Boya Wu, Qicong Chen, Juanzi Li, Chunxiao Xing, and Jie Tang. **How Do Your Friends on Social Media Disclose Your Emotions?** AAAI, 2014.

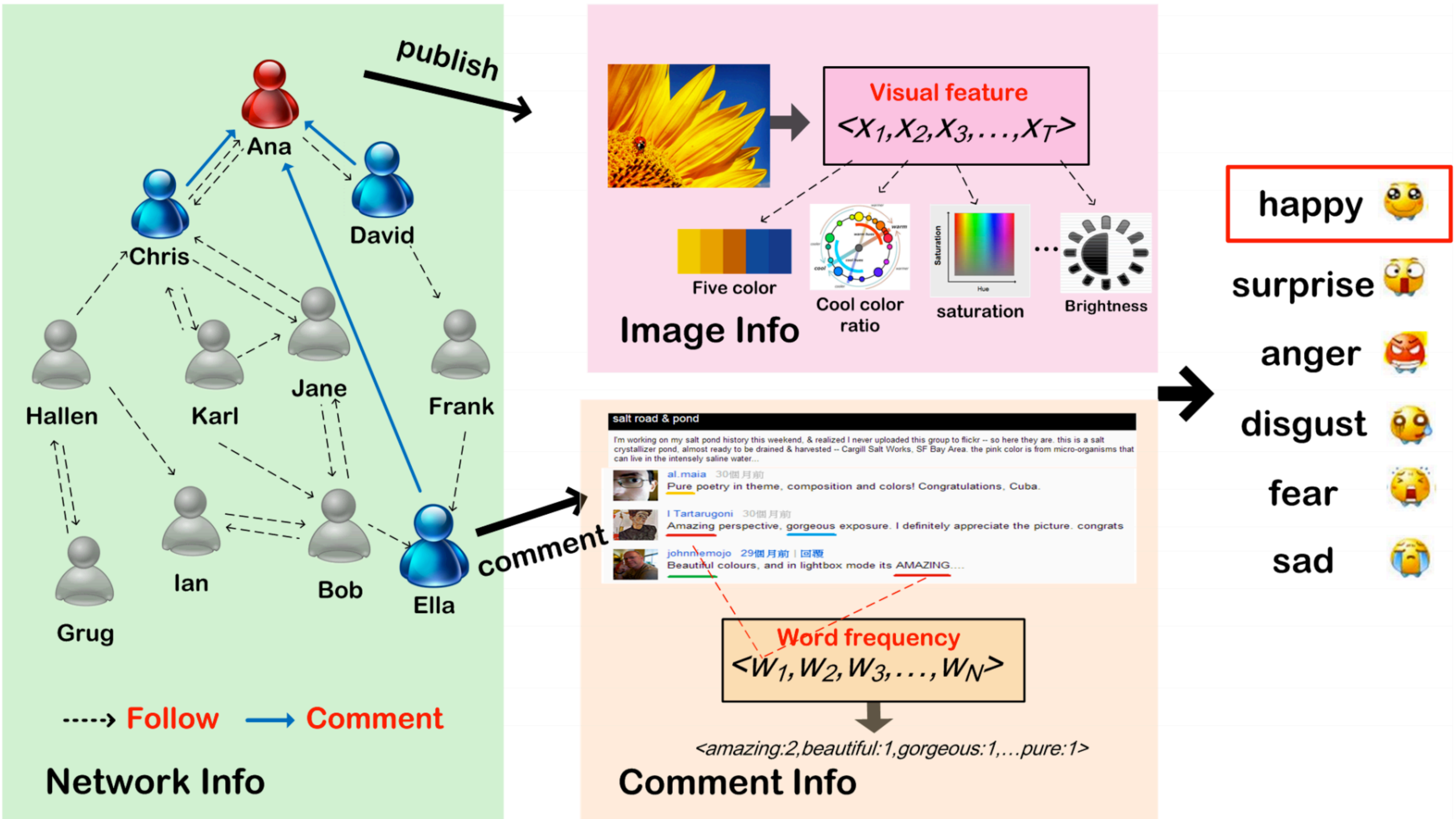
Was Anna Happy When She Published This Photo On Flickr?



To What Extent Your Friends Will Disclose Your Emotions?



Problem



Yang Yang, Jia Jia, Shumei Zhang, Boya Wu, Qicong Chen, Juanzi Li, Chunxiao Xing, and Jie Tang. **How Do Your Friends on Social Media Disclose Your Emotions?** AAAI, 2014.

Predicting Users' Emotional Status

- **Input:** An image social network G=<V, M, D, E, R, L>, where V is a set of users, M is a set of images, D is a set of comments, E represents following relationships between users, each element in R (v, m, t) denotes that user v publishes image m at time t, and an edge in L (v, d, m) indicates that user v leaves a comment d under image m.

- We use a matrix Y to denote users' emotional status, where y_{vt} indicates v's emotion at time t. $y_{vt} \in$ {happiness, surprise, anger, disgust, fear, sadness}

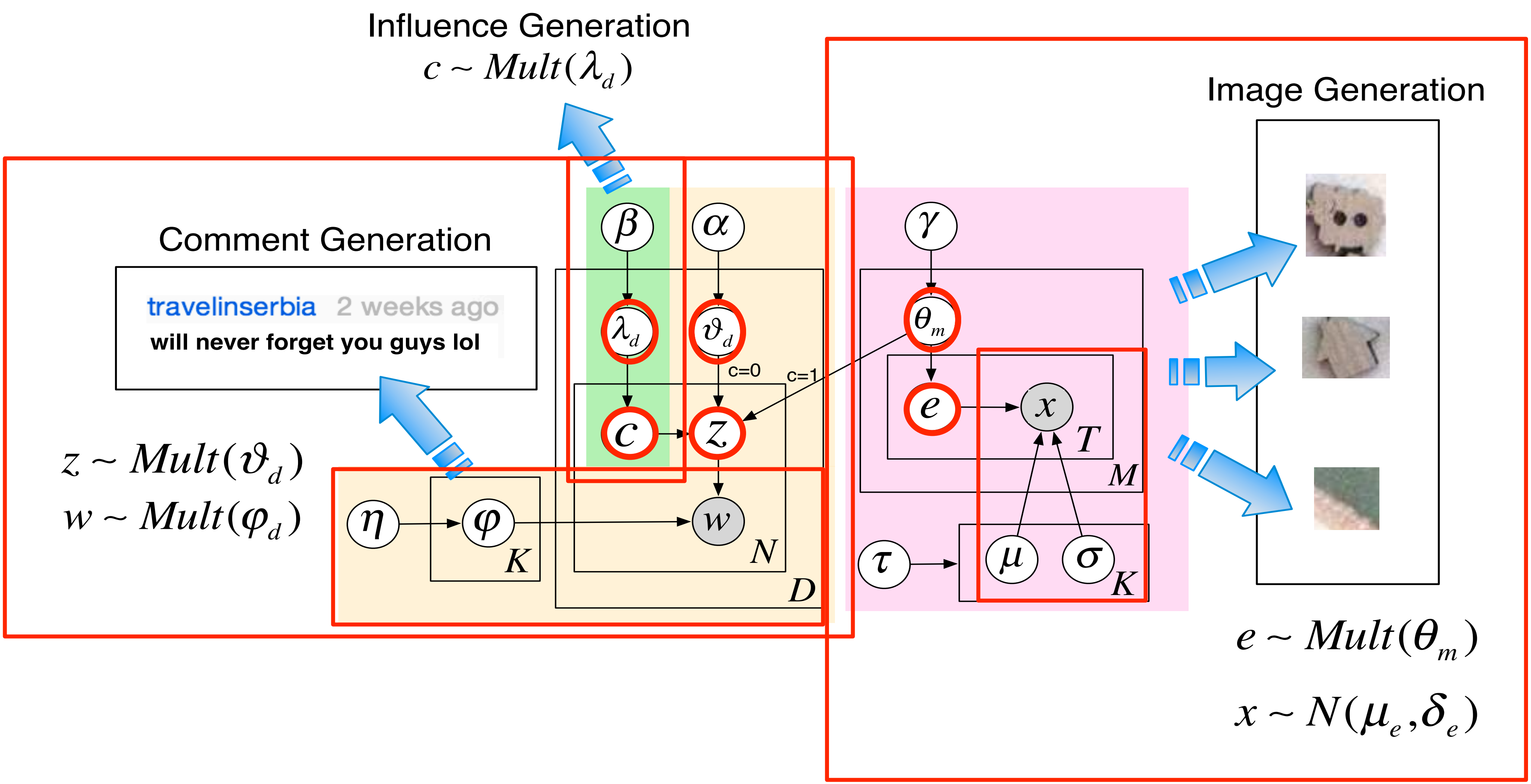
- Task: Given G, Y, a time stamp t, our goal is to learn

$$f : G = (V, M, E, R), t, Y_{.1\ldots t-1} \rightarrow Y_{.t}$$

Challenges

- How to model the image information and content information jointly?
- How to learn the association between the implied emotions of different comments?

Emotion Learning Method



Yang Yang, Jia Jia, Shumei Zhang, Boya Wu, Qicong Chen, Juanzi Li, Chunxiao Xing, and Jie Tang. **How Do Your Friends on Social Media Disclose Your Emotions?** AAAI, 2014.

Generative Process

Input: the hyper-parameters α, β_0, b_0, b_1, γ, η, and τ, the image-based social network G

foreach *image* $m \in M$ **do**
- **foreach** *visual feature* x_{mt} *of* m **do**
  - Generate $e_{mt} \sim \text{Mult}(\theta_m)$;
  - Generate $x_{mt} \sim N(x_{mt} | \mu_{e_{mt}t}, \delta_{e_{mt}t})$;
- **end**
- **foreach** *comment* d, *where* $a_{md} \in A$ **do**
  - **foreach** *word* w_{di} *of* d **do**
    - Generate $c_{di} \sim \text{Mult}(\lambda_d)$;
    - **if** $c_{di} == 0$ **then**
      - Generate $z_{di} \sim \text{Mult}(\theta_d)$;
    - **end**
    - **if** $c_{di} == 1$ **then**
      - Generate $z_{di} \sim \text{Mult}(\theta_m)$;
    - **end**
    - Generate $w_{di} \sim \text{Mult}(\varphi_{z_{di}})$
  - **end**
- **end**

end

Algorithm 1: Probabilistic generative process in the proposed model.

Visual feature generation

User influence generation

User comment generation

Learning Algorithm

- We employ Gibbs sampling to estimate unknown parameters.
  - The posterior for sampling the latent variables for each word:

$$P(z_{di}, c_{di} = 0 | \mathbf{z}_{\neg di}, \mathbf{c}_{\neg di}, \mathbf{w}) = \frac{n_{z_{di}d}^{\neg di} + \alpha}{\sum_z (n_{zd}^{\neg di} + \alpha)} \times \frac{n_{c_{di}d}^{\neg di} + \beta_{c_{di}}}{\sum_c (n_{cd}^{\neg di} + \beta_c)} \times \frac{n_{z_{di}w_{di}}^{\neg di} + \eta}{\sum_w (n_{z_{di}w}^{\neg di} + \eta)}$$

#(c_{di} is sampled associated with i-th word in d)

  - The posterior for sampling the latent emotion:

$$P(e_{mt}; \mathbf{e}_{\neg mt}, \mathbf{x}) = \frac{n_{me_{mt}}^{\neg mt} + \gamma}{\sum_e (n_{me}^{\neg mt} + \gamma)} \times \frac{\Gamma(\tau_2 + \frac{n_{e_{mt}t}}{2})}{\Gamma(\tau_2 + \frac{n_{e_{mt}t}^{\neg mt}}{2})} \times$$

$$\frac{\sqrt{\tau_1 + n_{e_{mt}t}^{\neg mt}} [\tau_3 + \frac{1}{2}(n_{e_{mt}t}^{\neg mt} s_{e_{mt}t}^{\neg mt} + \frac{\tau_1 n_{e_{mt}t}^{\neg mt} (\overline{x}_{e_{mt}t}^{\neg mt} - \tau_0)^2}{\tau_1 + n_{e_{mt}t}^{\neg mt}})]^{(\tau_2 + \frac{n_{e_{mt}t}^{\neg mt}}{2})}}{\sqrt{\tau_1 + n_{e_{mt}t}} [\tau_3 + \frac{1}{2}(n_{e_{mi}t} s_{e_{mt}t} + \frac{\tau_1 n_{e_{mt}t} (\overline{x}_{e_{mt}t} - \tau_0)^2}{\tau_1 + n_{e_{mt}t}})]^{(\tau_2 + \frac{n_{e_{mt}t}}{2})}}$$

use Stirling's formula to calculate gamma function

Learning Algorithm (cont.)

- Update for parameters of topic modeling part:

$$\theta_{dz} = \frac{n_{zd} + \alpha}{\sum_{z'}(n_{z'd} + \alpha)} \qquad \theta_{me} = \frac{n_{zm} + \gamma}{\sum_{e'}(n_{e'm} + \gamma)}$$

$$\lambda_{dc} = \frac{n_{cd} + \beta_c}{\sum_{c'} n_{c'd} + \beta_{c'}} \qquad \varphi_{zw} = \frac{n_{zw} + \eta}{\sum_{w'}(n_{zw'} + \eta)}$$

- The update for Gaussian parameters are hard to compute. We approximate Gaussian parameters by their expectations.

$$\mu_{et} \approx E(\mu_{et}) = \frac{\tau_0 \tau_1 + n_{et}\bar{x}_{et}}{\tau_1 + n_{et}}$$

$$\delta_{et} \approx E(\delta_{et}) = \frac{2\tau_2 + n_{et}}{2\tau_3 + n_{et}s_{et} + \frac{\tau_1 n_{et}(\bar{x}_{et} - \tau_0)^2}{\tau_1 + n_{et}}}$$

Flickr Data

- 354,192 images posted by 4,807 users
  - For each image, we also collect its tags and all comments.
  - We get 557,177 comments posted by 6,735 users in total
- Infer emotion of users by considering both image and tag/comments

Emotion Inference

Averagely *+37.4%* in terms of F1

Table 2: Performance of emotion inference.

Emotion	Method	Precision	Recall	F1-score	Emotion	Method	Precision	Recall	F1-score
Happiness	SVM	0.242	0.279	0.259	Disgust	SVM	0.192	0.236	0.212
	PFG	0.337	0.312	0.324		PFG	0.309	0.374	0.339
	LDA+SVM	0.333	**0.727**	**0.457**		LDA+SVM	0.223	0.223	0.223
	EL+SVM	**0.367**	0.410	0.388		EL+SVM	**0.331**	**0.432**	**0.374**
Surprise	SVM	0.197	0.037	0.063	Fear	SVM	0.204	0.264	0.230
	PFG	0.349	0.340	0.345		PFG	0.301	**0.408**	0.347
	LDA+SVM	0.218	0.048	0.078		LDA+SVM	0.211	0.225	0.217
	EL+SVM	**0.425**	**0.516**	**0.466**		EL+SVM	**0.371**	0.343	**0.356**
Anger	SVM	0.188	0.105	0.135	Sadness	SVM	0.225	0.365	0.278
	PFG	0.191	0.142	0.163		PFG	0.357	0.286	0.317
	LDA+SVM	0.222	0.109	0.146		LDA+SVM	0.257	0.278	0.267
	EL+SVM	**0.390**	**0.370**	**0.380**		EL+SVM	**0.561**	**0.617**	**0.588**

SVM: regards the visual features of images as inputs and uses a SVM as a classifier.

PFG: considers both color features and social correlations among images.

LDA+SVM: first uses LDA to extract latent topics from comments, then uses visual features, topic distributions, and social ties as features to train a SVM.

To What Extend Your Friends Can Disclose Your Emotions?

-Comments stands for the proposed method ignoring comment information

-Tie ignores social tie information

Fear images have similar visual features with Sadness and Anger.

Homophily suggests that friends with similar interests tend to have similar understanding of disgust

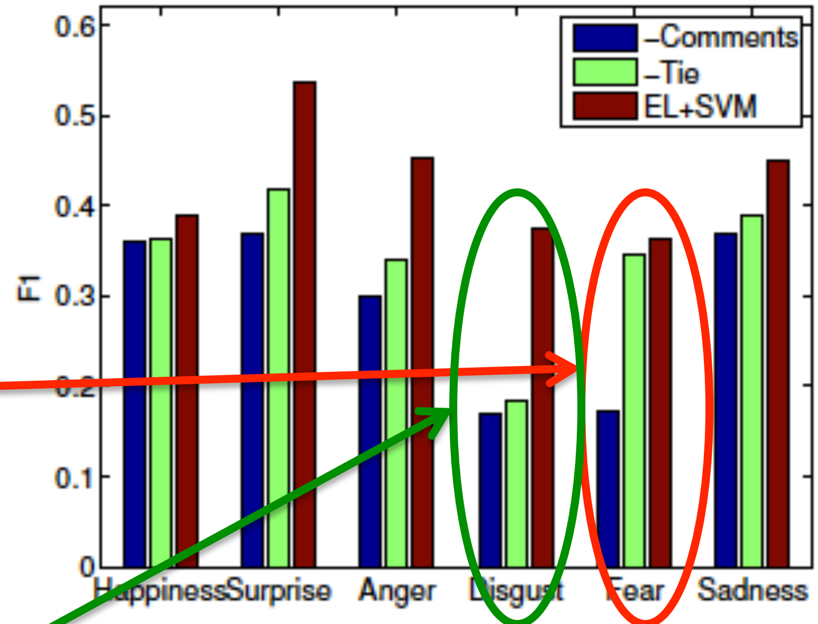
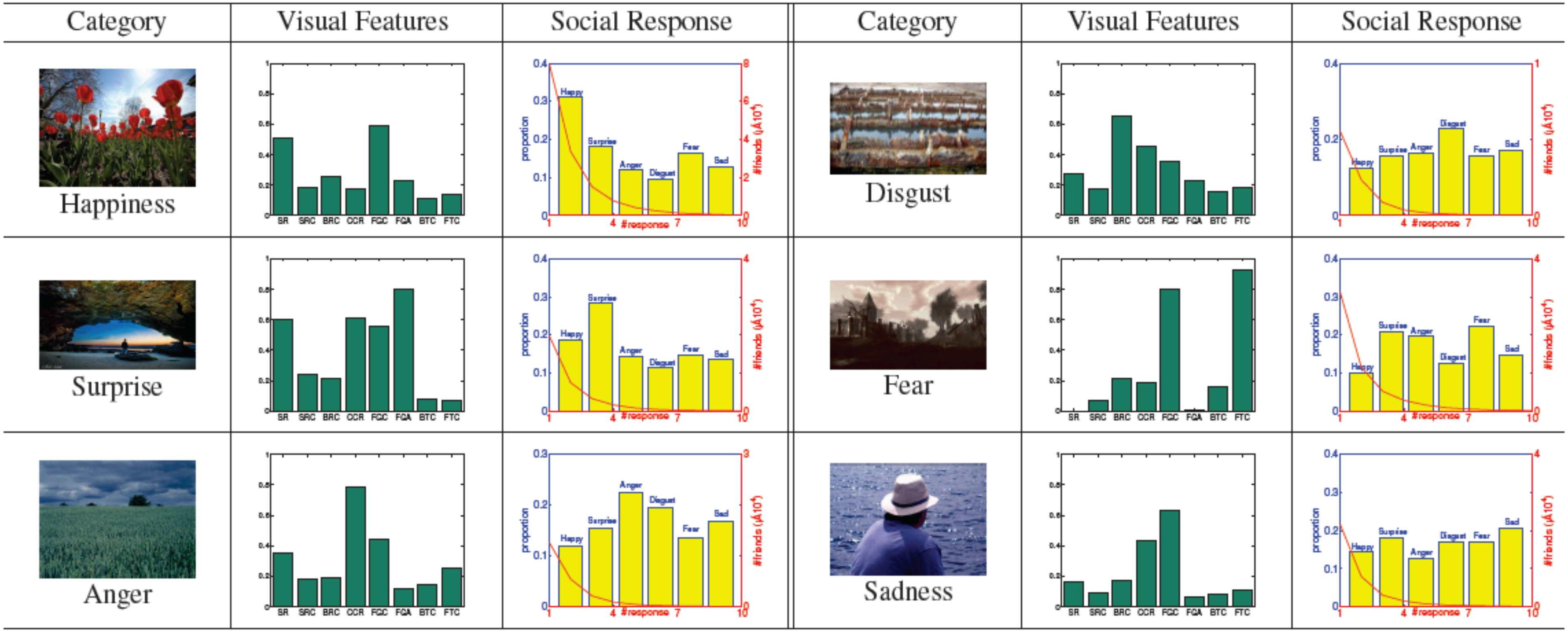


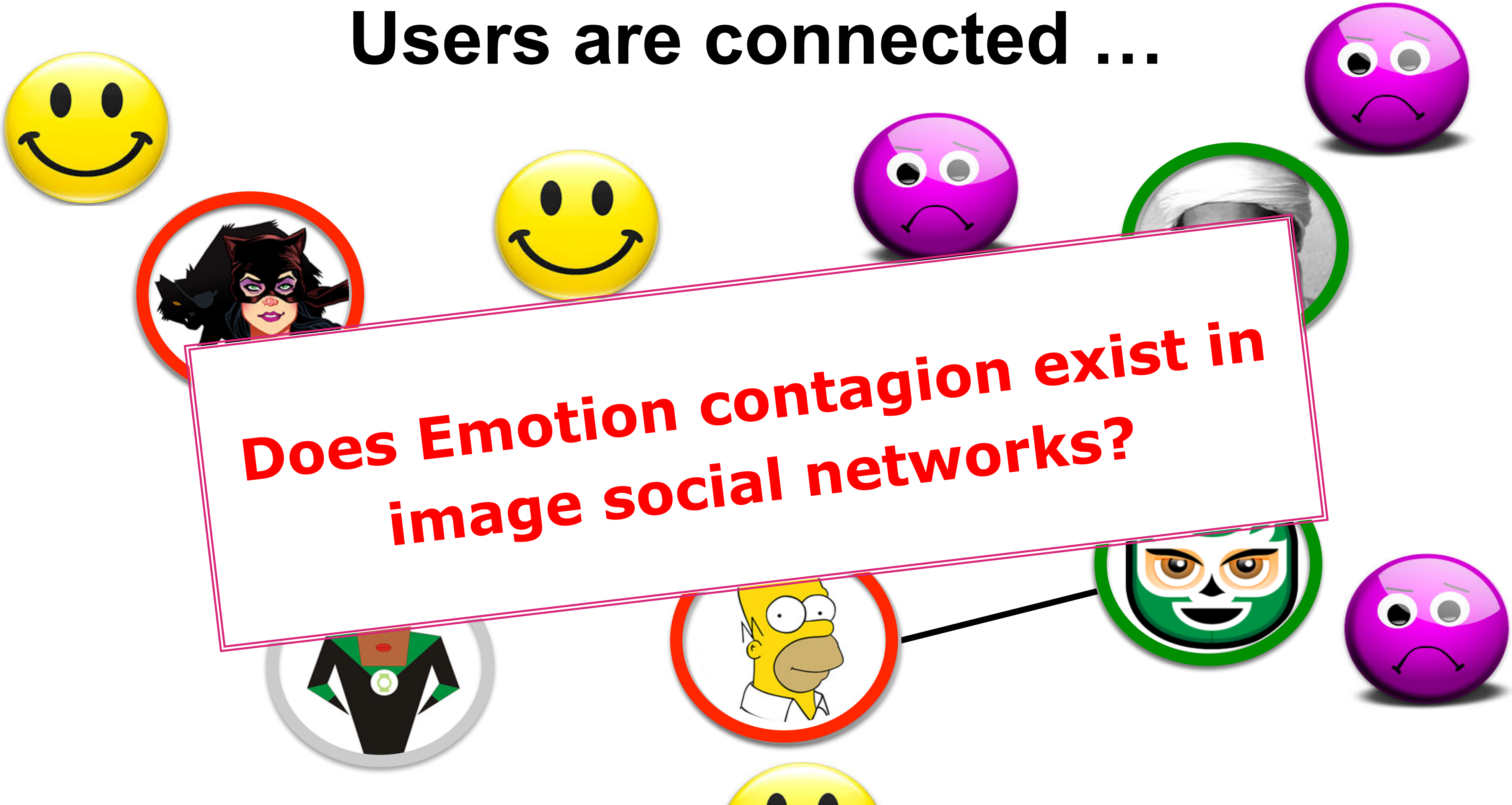
Image Interpretations



- Our model demonstrates how visual features distribute over different emotions. (e.g., images representing Happiness have high saturation)
- Positive emotions attract more response (***+4.4*** times) and more easily to influence others compared with negative emotions.

What will Happen after Spiderman Posts this Photo?

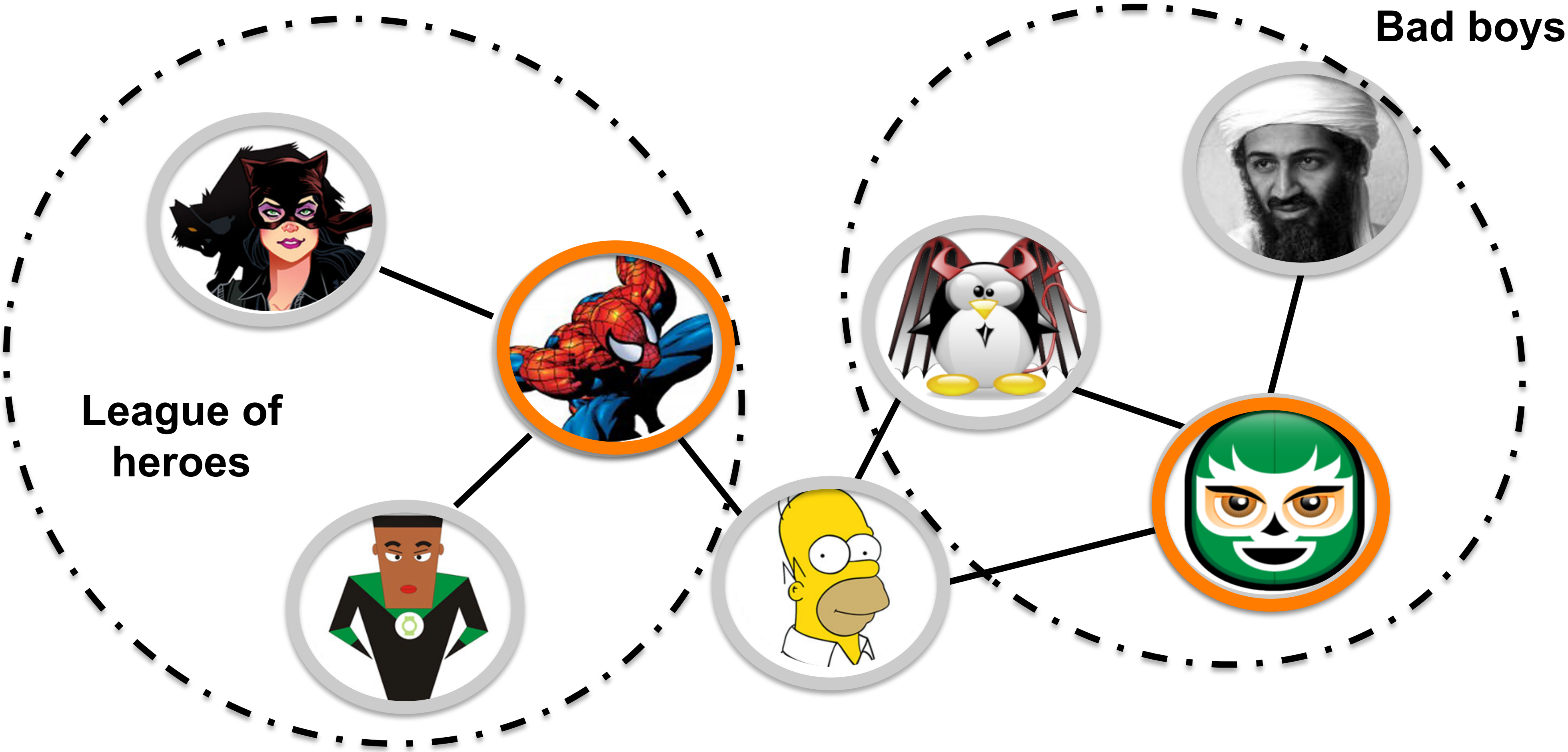
Users are connected …



Does Emotion contagion exist in image social networks?

Emotion Contagion: The cascade of users' emotional statuses influence each other

Social Roles of Users



Opinion leaders: users taking central positions in communities

Social Roles of Users

Will social roles influence emotion contagions?

Structural hole spanners: users bridge otherwise disconnected communities

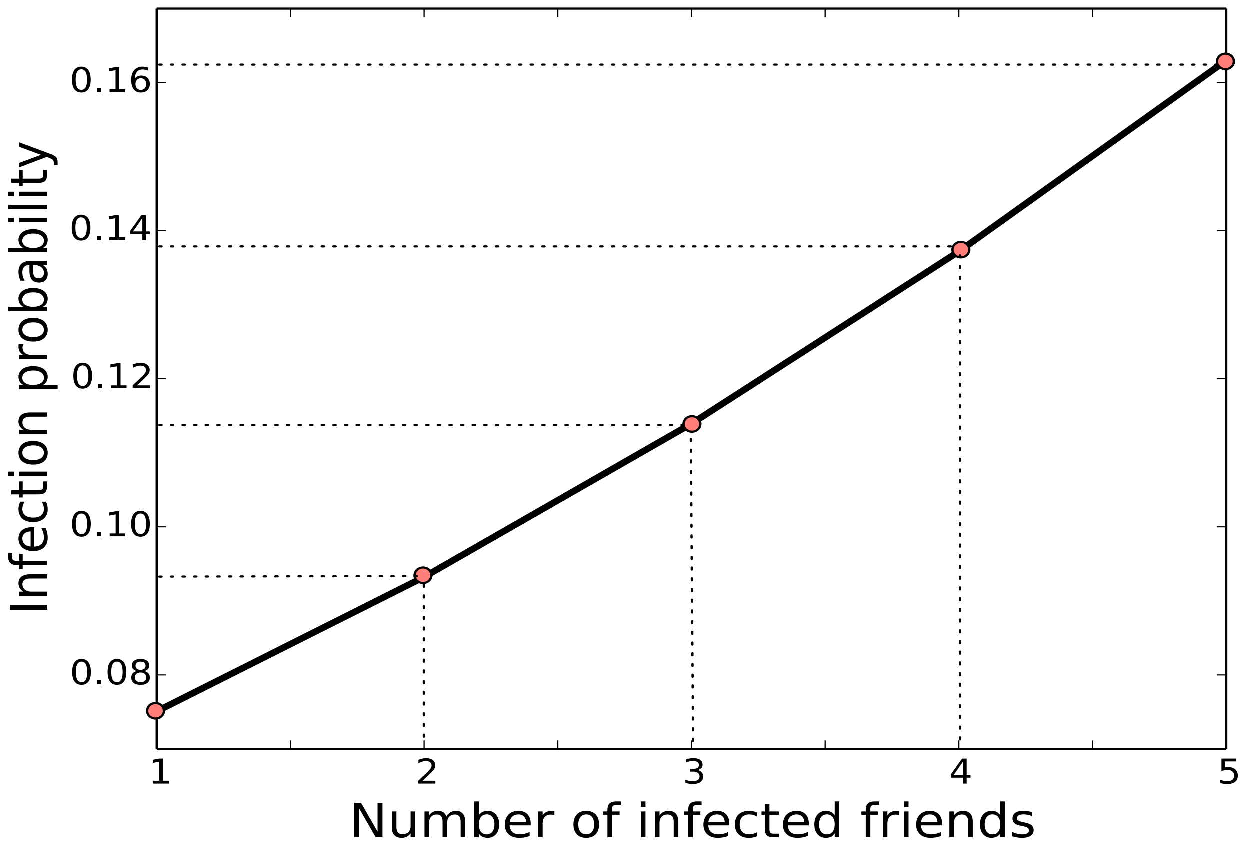
Three Qs to Answer

- **Q1:** Does emotion contagion exist in image social networks?
- **Q2:** Will social roles influence emotion contagion?
- **Q3:** How to better predict the emotional status of users in social networks by considering emotion contagion?

Yang Yang, Jia Jia, Boya Wu, and Jie Tang. **Social Role-Aware Emotion Contagion in Image Social Networks.** AAAI, 2016.

Q1: Existence

Q1.1: When your friends are happy, will you be happy?



Yang Yang, Jia Jia, Boya Wu, and Jie Tang. **Social Role-Aware Emotion Contagion in Image Social Networks.** AAAI, 2016.

Q1: Existence

Q1.2: When predicting a user's emotional status, will her friends help?

Historical post logs
+
Previous emotion
+
Image features

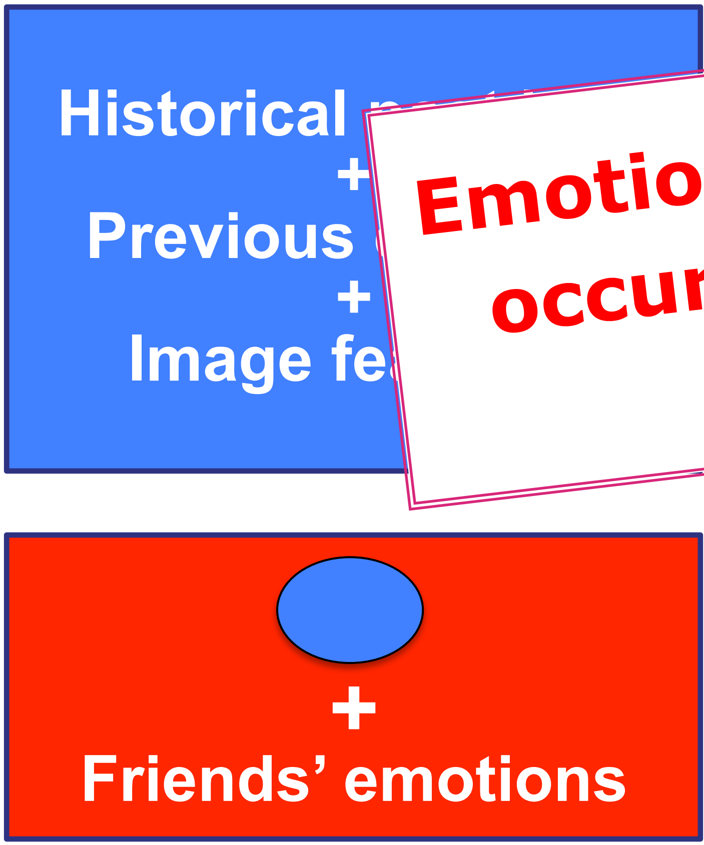
+
Friends' emotions

Predict →

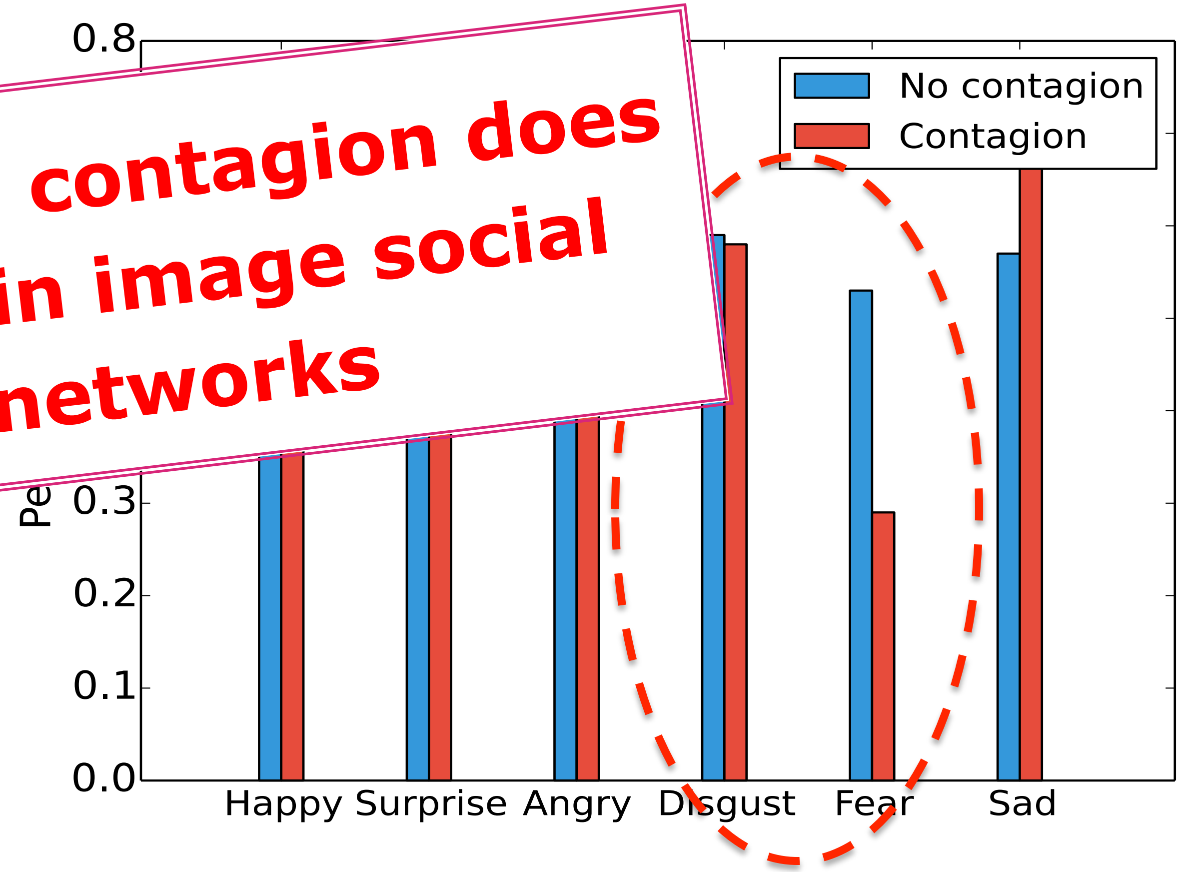
Y_{vt}
User v's emotional status at time t
happiness, surprise, anger, disgust, fear, sadness

Q1: Existence

Q1.2: When predicting a user's emotional status, will her friends help?



Emotion contagion does occur in image social networks



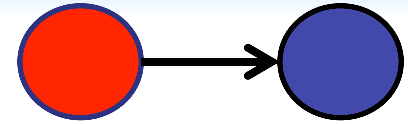
Q2: Social Role

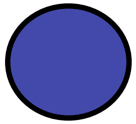
- ***Opinion leaders:*** 20% of users with largest PageRank scores;
- ***Structural hole spanners:*** 20% of users with lowest network constraint scores;
- Others are remaining as ***ordinary users***.

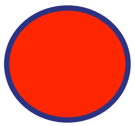
OL and SH y users in

Still holds in emotion contagion?

Q2: Social Role



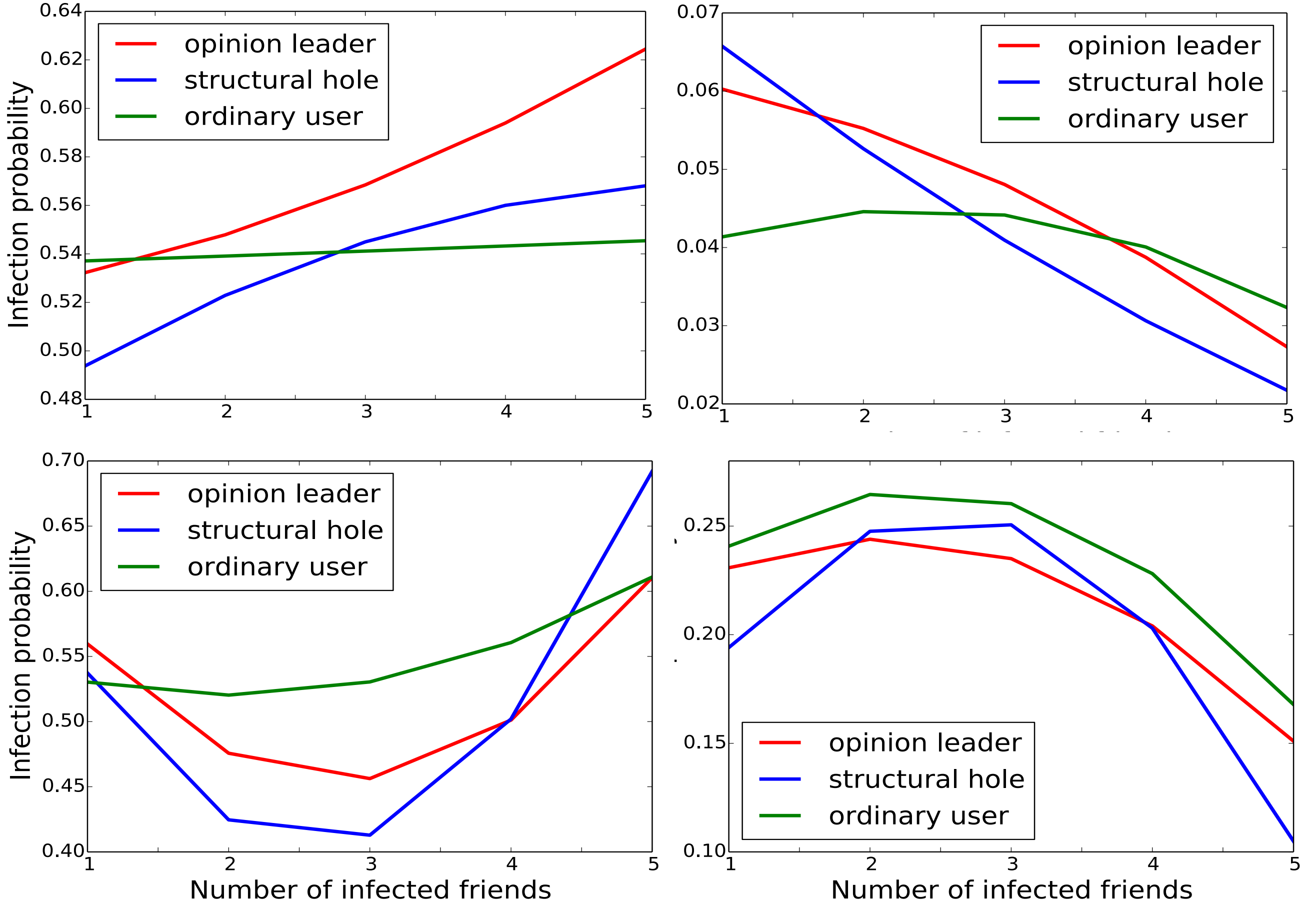




Happy **Fear**

Happy

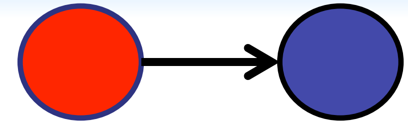
Fear

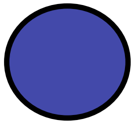


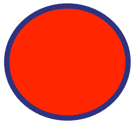
X: number of friends with different social roles.

Y: probability being a certain emotion.

Q2: Social Role



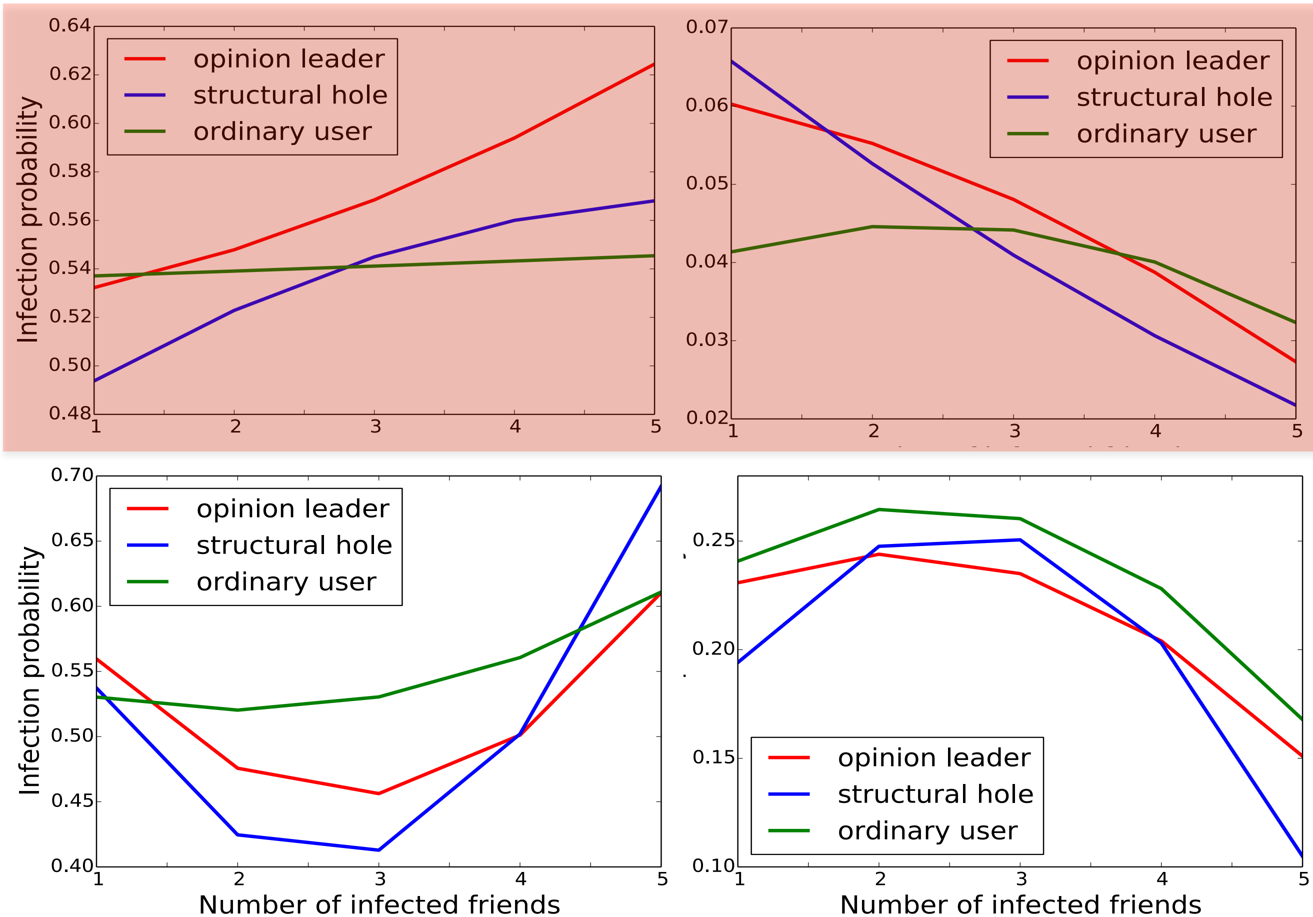




Happy **Fear**

Happy

Fear

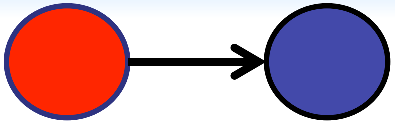


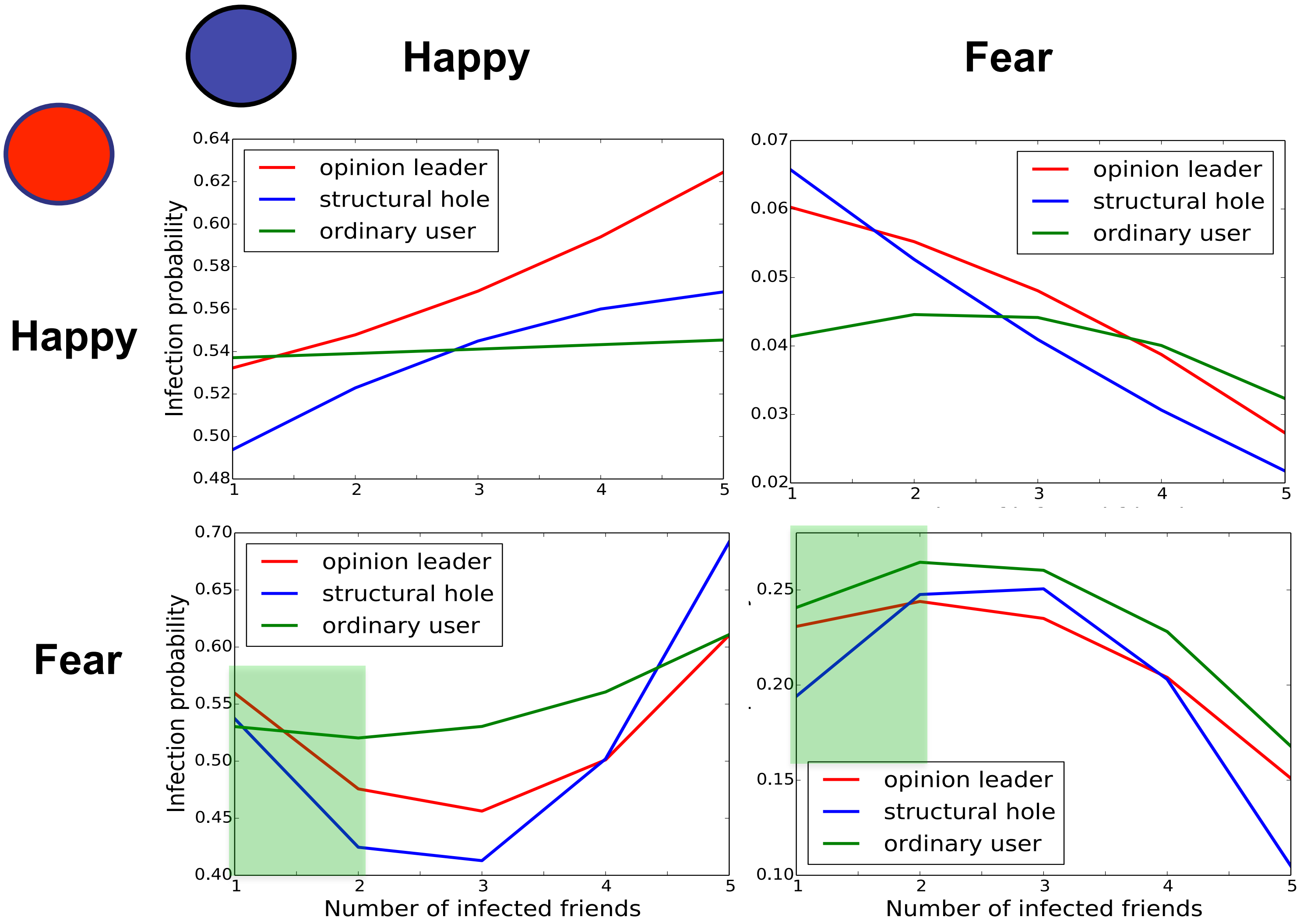
X: number of friends with different social roles.

Y: probability being a certain emotion.

positive emotion delights friends

Q2: Social Role

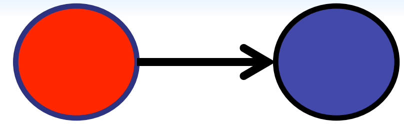


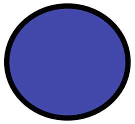


X: number of friends with different social roles.

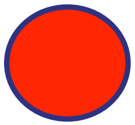
Y: probability being a certain emotion.

Q2: Social Role



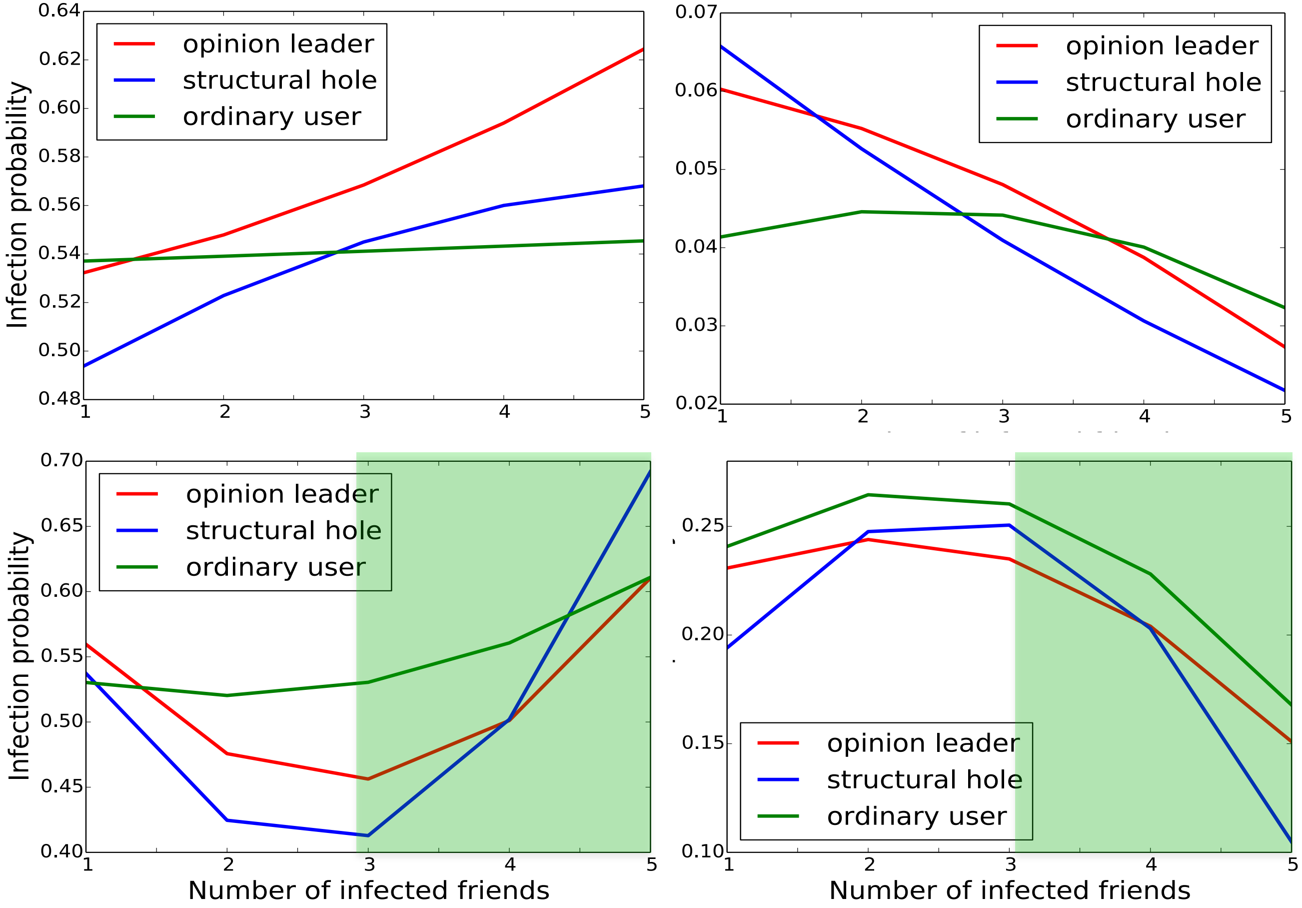


Happy | **Fear**



Happy

Fear

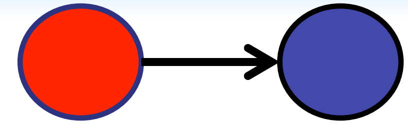


X: number of friends with different social roles.

Y: probability being a certain emotion.

"Emotional comfort" phenomena

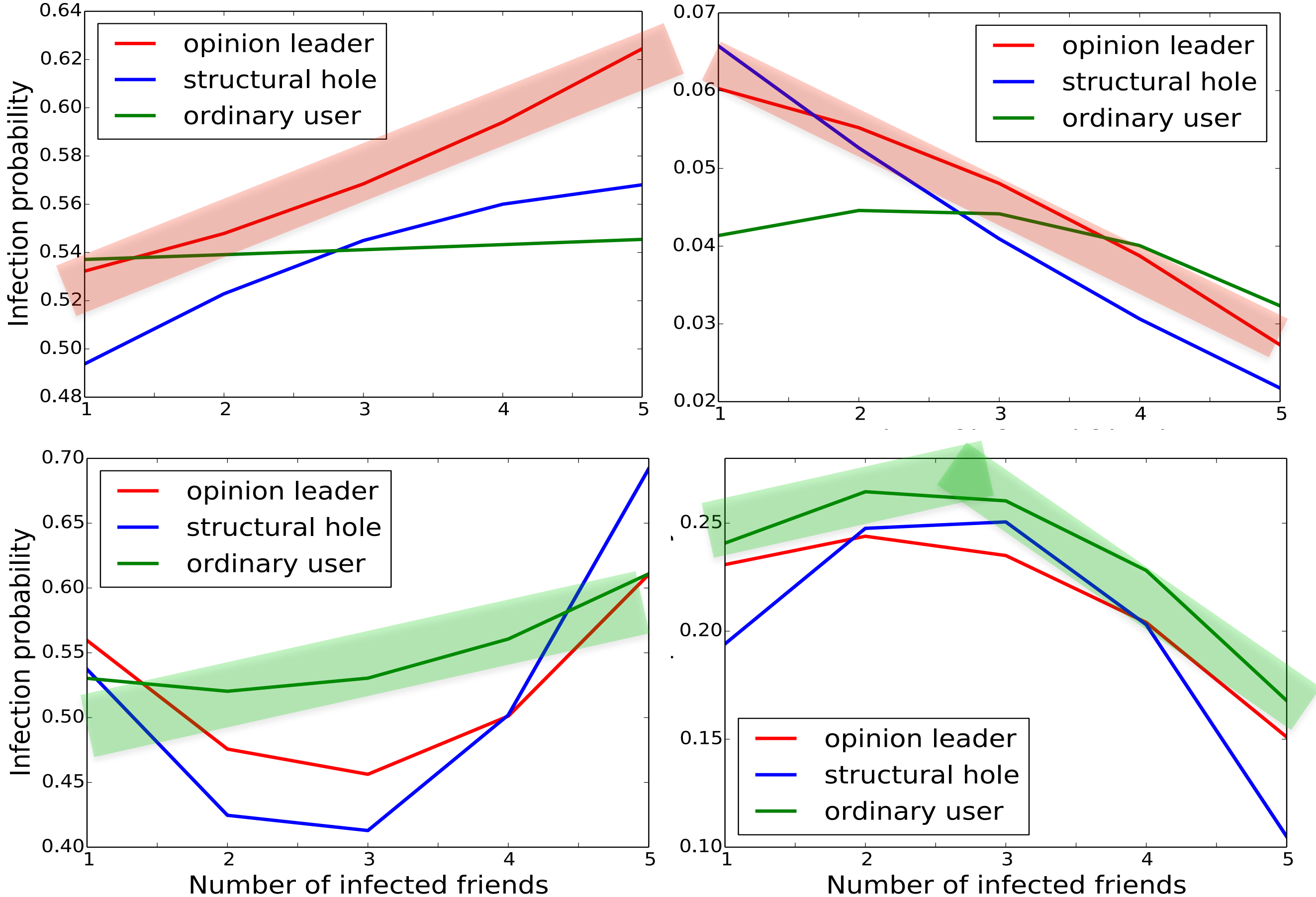
Q2: Social Role



Happy | Fear

Happy

Fear



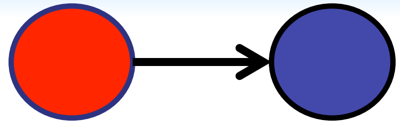
X: number of friends with different social roles.

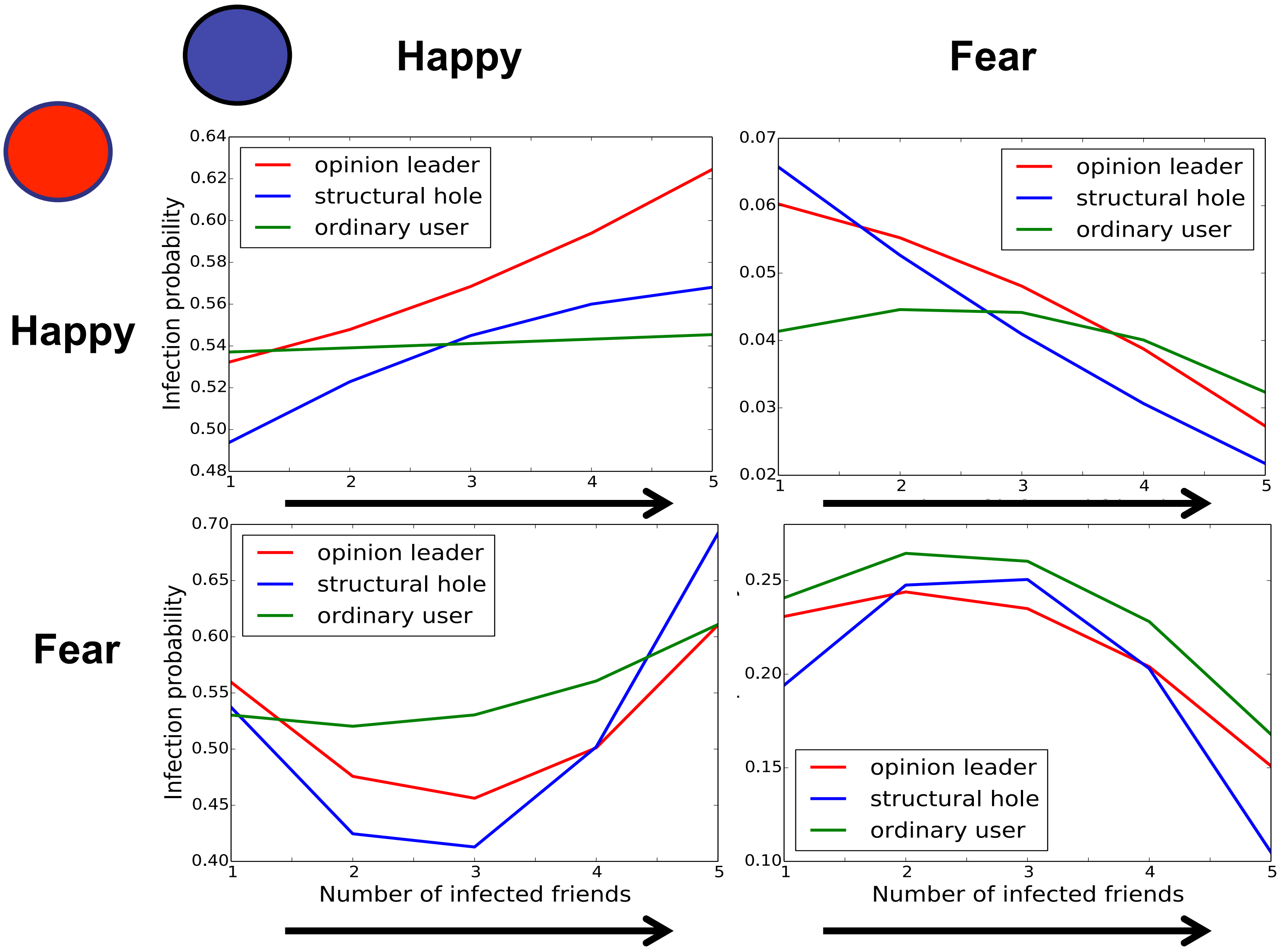
Y: probability being a certain emotion.

Opinion leaders are more influential on positive emotions

Ordinary users are more influential on negative emotions

Q2: Social Role



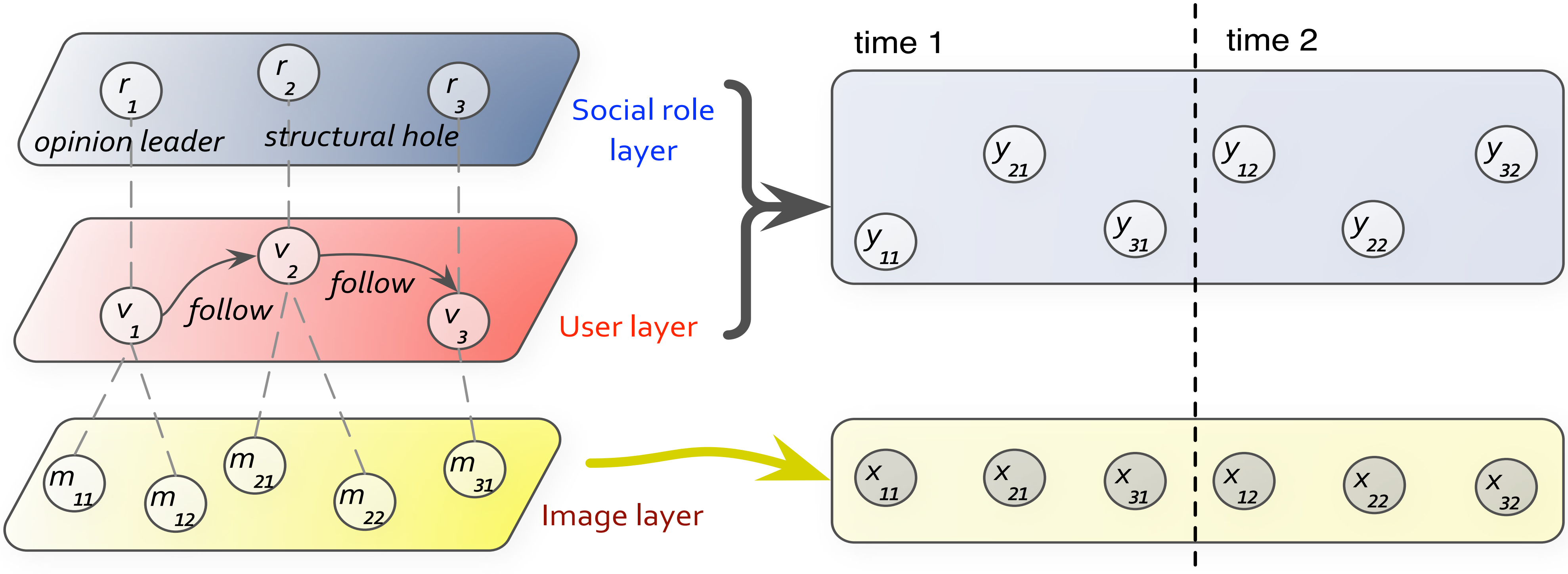


X: number of friends with different social roles.

Y: probability being a certain emotion.

Influence of opinion leaders and structural holes change faster than ordinary users.

Q3: Model

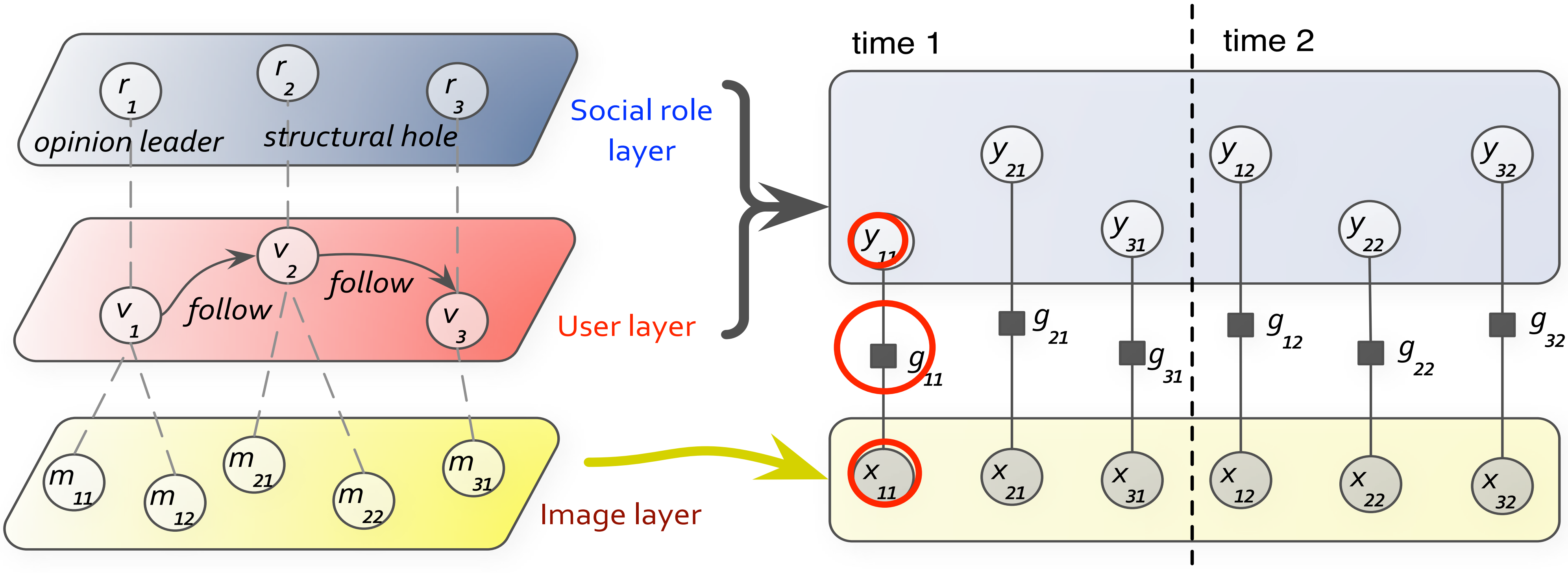


(a) An example of the problem (b) Social Role-Aware Contagion Model

P(Y|G): Conditional probability of users' emotional status given input data

Q3: Model

P(Y|G)=πg(.) …



(a) An example of the problem

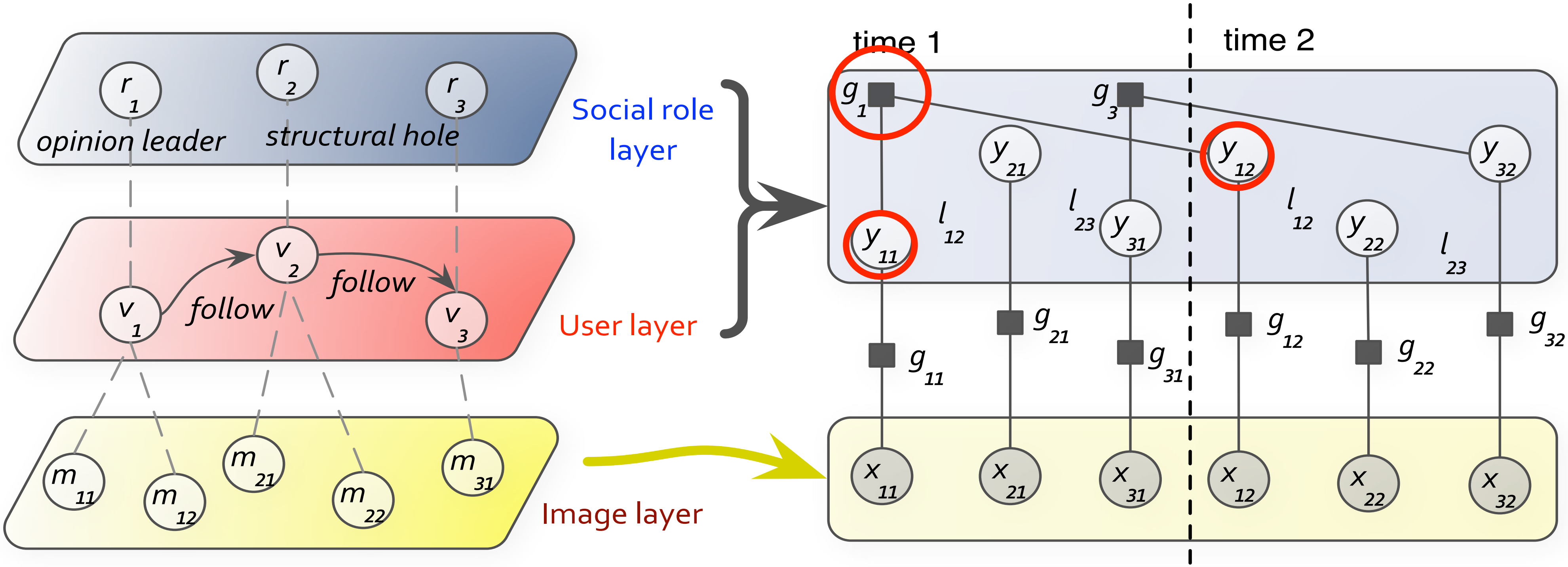
(b) Social Role-Aware Contagion Model

$g(x_{vt}, y_{vt})$: Correlation between v's emotion and the image she posts at t.

$$g(x_{vt}, y_{vt}) = \frac{1}{Z_1} \exp\{\alpha_{y_{vt}} \cdot x_{vt}\}$$

Q3: Model

P(Y|G)=π{g(.)h(.)} …



(a) An example of the problem

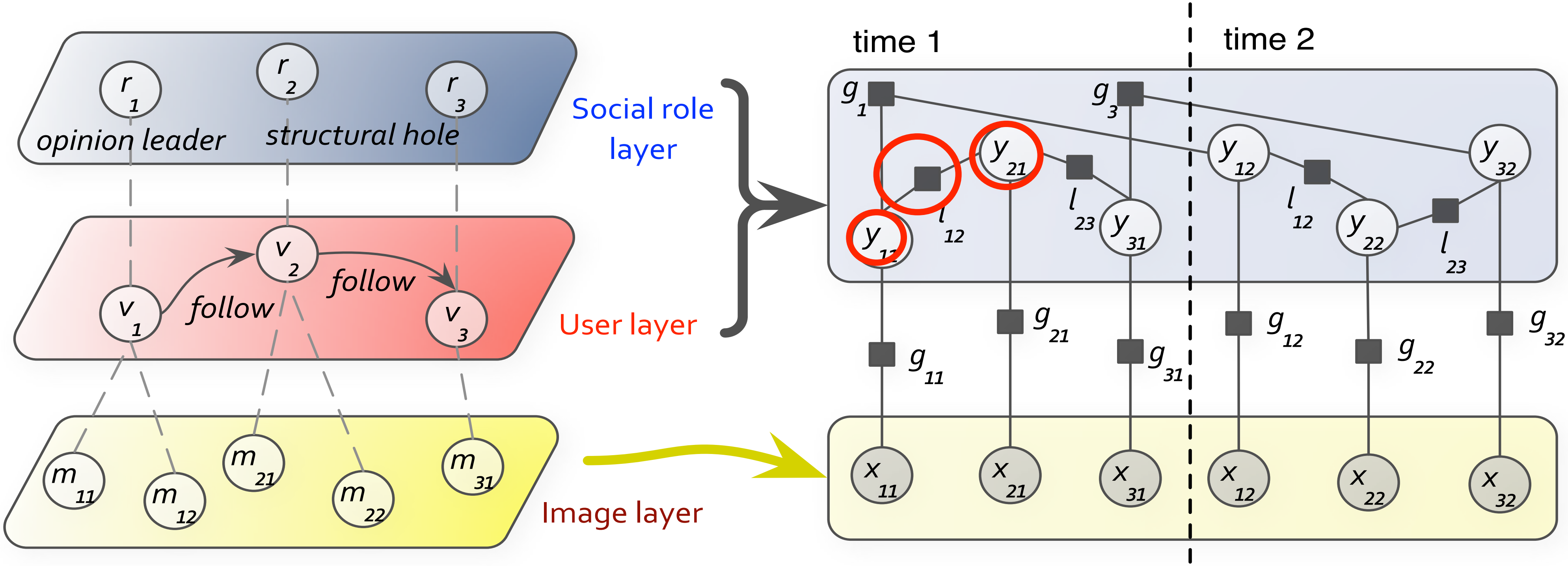
(b) Social Role-Aware Contagion Model

h($y_{ut-t'}$, y_{vt}): Correlation between v's emotion at time t and t-t'.

$$h(y_{vt-\Delta t}, y_{vt}) = \frac{1}{Z_2} \exp\{\beta_{\Delta t} \cdot I(y_{vt-\Delta t}, y_{vt})\}$$

Q3: Model

P(Y|G)=π{g(.)h(.)l(.)}



(a) An example of the problem

(b) Social Role-Aware Contagion Model

$l(y_{ut-1}, y_{vt})$: How v's emotion at t is influenced by her friend u's emotion at t-1.

$$l(y_{ut-1}, y_{vt}) = \frac{1}{Z_3} \exp\{\boxed{\gamma_{r_u r_v}} \cdot I(y_{ut-1}, y_{vt})\}$$

Social role sensitive parameter

Experimental Results

Emotion
Happiness
Surprise
Anger

Flickr dataset:

2,060,353 images, 1,255,478 users
ground truth obtained by user tags

Distribution of users' emotional statuses on Flickr:

happiness: 46.2%
surprise: 9.7%
anger: 8.0%
disgust: 5.3%
fear:17.3%
sadness: 13.5%

Experimental Results

Emotion	Method
Happiness	SVM
	LR
	NB
	BN
	RBF
	CRF
	Role-aware
Surprise	SVM
	LR
	NB
	BN
	RBF
	CRF
	Role-aware
Anger	SVM
	LR
	NB
	BN
	RBF
	CRF
	Role-aware

Baselines

Methods do not consider emotion contagion:
SVM, Logistic Regression (LR),
Naïve Bayes (NB), Bayesian Network (BN),
Gaussian Radial Basis Function Neural Network (RBF).

Methods ignore social role information: CRF

Our model: Role-aware

Experimental Results

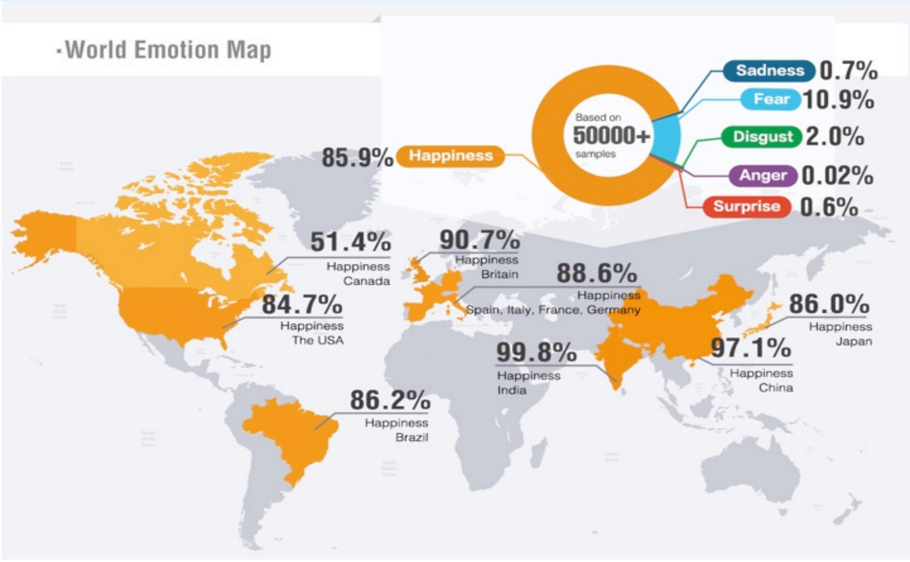
Emotion	Method	Precision	Recall	F1-score	Emotion	Method	Precision	Recall	F1-score
Happiness	SVM								
	LR								
	NB								
	BN								
	RBF								
	CRF								
	Role-aware								
Surprise	SVM								
	LR								
	NB								
	BN								
	RBF								
	CRF								
	Role-aware								
Anger	SVM								
	LR								
	NB								
	BN								
	RBF								
	CRF								
	Role-aware								

Evaluation Metrics:

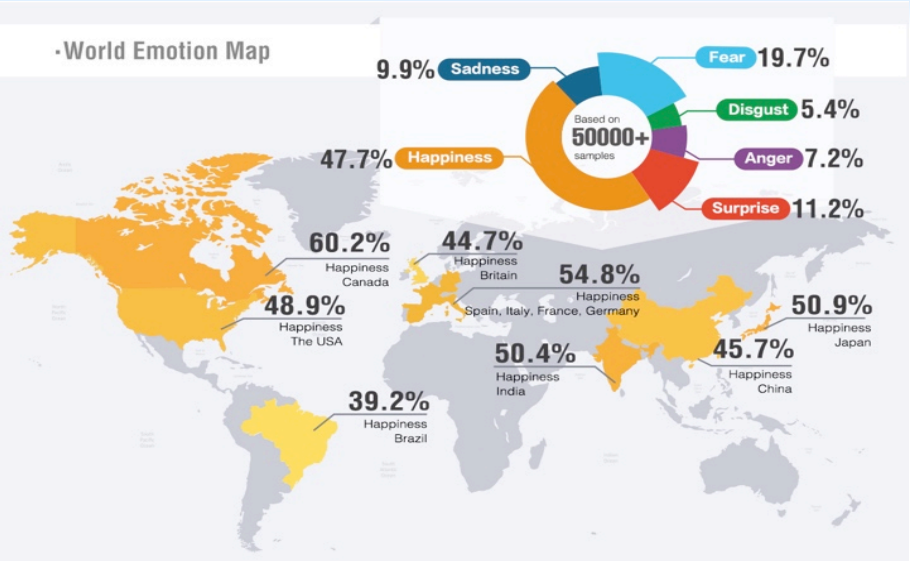
- Precision
- Recall
- F1 Measure

Experimental Results

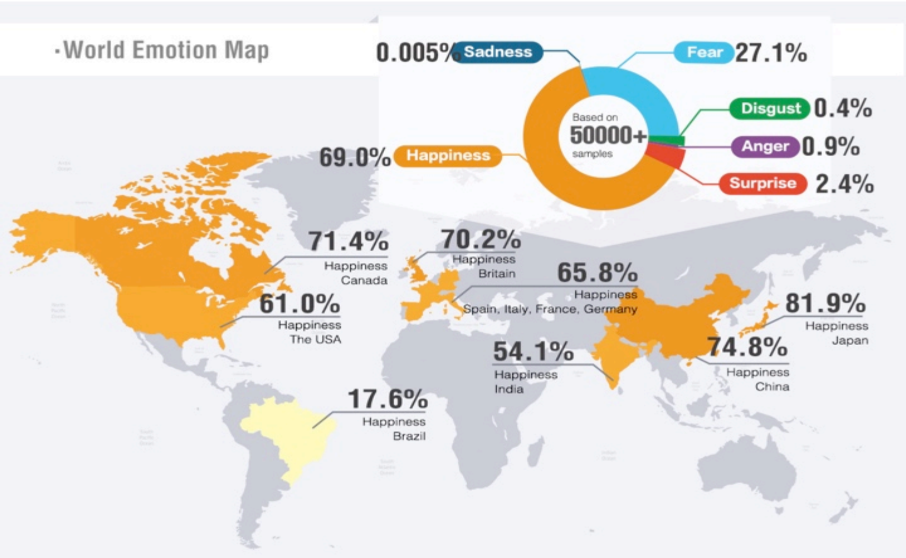
Emotion	Method	Precision	Recall	F1-score	Emotion	Method	Precision	Recall	F1-score
Happiness	SVM	0.5490	0.4682	0.5054	Disgust	SVM	0.5721	0.6223	0.5962
	LR	**0.5726**	0.4234	0.4868		LR	0.5902	0.5847	0.5874
	NB	0.5604	0.4679	0.5100		NB	0.5657	0.7244	0.6353
	BN	0.5605	0.5129	0.5357		BN	0.5666	0.6811	0.6186
	RBF	0.5744	0.2676	0.3651		RBF	0.5246	0.4346	0.4754
	CRF	0.5590	0.5938	0.5759		CRF	0.8304	0.5889	0.6891
	Role-aware	0.5285	**0.9327**	**0.6747**		Role-aware	**0.9758**	**0.9947**	**0.9852**
Surprise	SVM	0.5103	0.4821	0.4958	Fear	SVM	0.5253	0.5521	0.5384
	LR	0.5231	0.4108	0.4602		LR	0.5523	0.4703	0.5080
	NB	0.5124	0.5324	0.5222		NB	0.5350	0.5295	0.5322
	BN	0.5241	0.4712	0.4963		BN	0.5446	0.5189	0.5315
	RBF	0.4990	0.1756	0.2597		RBF	0.5227	0.2859	0.3696
	CRF	0.5810	0.8014	0.6736		CRF	0.5074	0.2123	0.2993
	Role-aware	**0.8992**	**0.9181**	**0.9086**		Role-aware	**0.8123**	**0.9996**	**0.8963**
Anger	SVM	0.5186	0.6371	0.5718	Sadness	SVM	0.5733	0.5740	0.5723
	LR	0.5275	0.4634	0.4934		LR	0.5664	0.4866	0.5234
	NB	0.5201	0.4959	0.5078		NB	0.5632	0.4991	0.5292
	BN	0.5260	0.5207	0.5233		BN	0.5730	0.5662	0.5695
	RBF	0.5062	0.2441	0.3294		RBF	0.5344	0.4292	0.4761
	CRF	0.6036	0.8015	0.6886		CRF	0.6382	0.8726	0.7372
	Role-aware	**0.9346**	**0.9593**	**0.9468**		Role-aware	**0.8741**	**0.9550**	**0.9128**



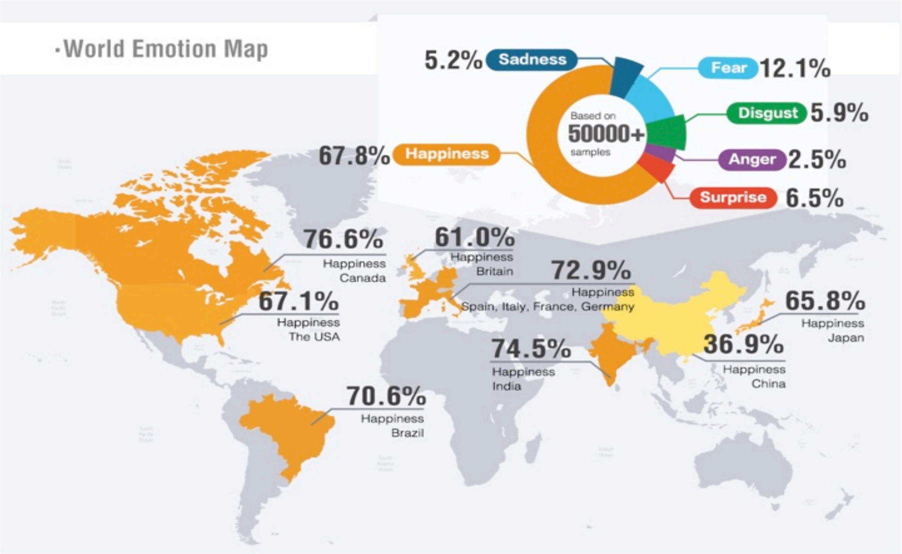
(a) Ground truth



(b) Random users



(c) Opinion leaders



(d) Structural hole spanners

Summary

- Learning social influence from multiple aspects
  - **Topic-based** social influence learning
  - **Social role-aware** influence learning
- Application: How user **emotions** diffuse in social networks
- Current work
  - Social influence based **representation learning** for dynamic networks

Related Publications

- Jie Tang, Jimeng Sun, Chi Wang, and Zi Yang. Social Influence Analysis in Large-scale Networks. In **KDD'09**, pages 807-816, 2009.
- Jie Tang, Jing Zhang, Limin Yao, Juanzi Li, Li Zhang, and Zhong Su. ArnetMiner: Extraction and Mining of Academic Social Networks. In **KDD'08**, pages 990-998, 2008.
- Yang Yang, Jia Jia, Boya Wu, and Jie Tang. Social Role-Aware Emotion Contagion in Image Social Networks. **AAAI'16**.
- Yang Yang, Jie Tang, Cane Wing-Ki Leung, Yizhou Sun, Qicong Chen, Juanzi Li, and Qiang Yang. RAIN: Social Role-Aware information Diffusion. **AAAI'15**.
- Yang Yang, Jia Jia, Shumei Zhang, Boya Wu, Qicong Chen, Juanzi Li, Chunxiao Xing, and Jie Tang. How Do Your Friends on Social Media Disclose Your Emotions? **AAAI'14**.
- Chenhao Tan, Jie Tang, Jimeng Sun, Quan Lin, and Fengjiao Wang. Social action tracking via noise tolerant time-varying factor graphs. In **KDD'10**, pages 807–816, 2010.
- Chenhao Tan, Lillian Lee, Jie Tang, Long Jiang, Ming Zhou, and Ping Li. User-level sentiment analysis incorporating social networks. In **KDD'11**, pages 1397–1405, 2011.
- Jia Jia, Sen Wu, Xiaohui Wang, Peiyun Hu, Lianhong Cai, and Jie Tang. Can We Understand van Gogh's Mood? Learning to Infer Affects from Images in Social Networks. In **ACM MM**, pages 857-860, 2012.
- Lu Liu, Jie Tang, Jiawei Han, Meng Jiang, and Shiqiang Yang. Mining Topic-Level Influence in Heterogeneous Networks. In **CIKM'10**, pages 199-208, 2010.
- Tiancheng Lou, Jie Tang, John Hopcroft, Zhanpeng Fang, Xiaowen Ding. Learning to Predict Reciprocity and Triadic Closure in Social Networks. In **TKDD**.
- Jimeng Sun and Jie Tang. Models and Algorithms for Social Influence Analysis. In **WSDM'13**. (Tutorial)
- Lu Liu, Jie Tang, Jiawei Han, and Shiqiang Yang. Learning Influence from Heterogeneous Social Networks. In **DMKD**, 2012, Volume 25, Issue 3, pages 511-544.
- Jimeng Sun and Jie Tang. A Survey of Models and Algorithms for Social Influence Analysis. Social Network Data Analytics, Aggarwal, C. C. (Ed.), Kluwer Academic Publishers, pages 177–214, 2011.
- J. Tang, S. Wu, and J. Sun. Confluence: Conformity Influence in Large Social Networks. In **KDD'13**.
- Jimeng Sun and Jie Tang. Models and Algorithms for Social Influence Analysis. In WSDM'13. (Tutorial)
- Chi Wang, Jie Tang, Jimeng Sun, and Jiawei Han. Dynamic Social Influence Analysis through Time-dependent Factor Graphs. In **ASONAM'11**, pages 239-246, 2011.
- Boya Wu, Jia Jia, Yang Yang, Peijun Zhao, and Jie Tang. Understanding the Emotions Behind Social Images: Inferring with User Demographics. ICME'15.

References

- S. Milgram. The Small World Problem. **Psychology Today**, 1967, Vol. 2, 60–67
- J.H. Fowler and N.A. Christakis. The Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis Over 20 Years in the Framingham Heart Study. **British Medical Journal** 2008; 337: a2338
- R. Dunbar. Neocortex size as a constraint on group size in primates. **Human Evolution**, 1992, 20: 469–493.
- R. M. Bond, C. J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle and J. H. Fowler. A 61-million-person experiment in social influence and political mobilization. **Nature**, 489:295-298, 2012.
- http://klout.com
- Why I Deleted My Klout Profile, by Pam Moore, at **Social Media Today**, originally published November 19, 2011; retrieved November 26 2011
- S. Aral and D Walker. Identifying Influential and Susceptible Members of Social Networks. **Science**, 337:337-341, 2012.
- J. Ugandera, L. Backstromb, C. Marlowb, and J. Kleinberg. Structural diversity in social contagion. **PNAS**, 109 (20):7591-7592, 2012.
- S. Aral, L. Muchnik, and A. Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. **PNAS**, 106 (51):21544-21549, 2009.
- J. Scripps, P.-N. Tan, and A.-H. Esfahanian. Measuring the effects of preprocessing decisions and network forces in dynamic network analysis. In **KDD’09**, pages 747–756, 2009.
- Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. **Journal of Educational Psychology** 66, 5, 688–701.
- http://en.wikipedia.org/wiki/Randomized_experiment

References(cont.)

- A. Anagnostopoulos, R. Kumar, M. Mahdian. Influence and correlation in social networks. In **KDD’08**, pages 7-15, 2008.
- L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report SIDL-WP-1999-0120, Stanford University, 1999.
- G. Jeh and J. Widom. Scaling personalized web search. In **WWW '03**, pages 271-279, 2003.
- G. Jeh and J. Widom, SimRank: a measure of structural-context similarity. In **KDD’02**, pages 538-543, 2002.
- A. Goyal, F. Bonchi, and L. V. Lakshmanan. Learning influence probabilities in social networks. In **WSDM’10**, pages 207–217, 2010.
- P. Domingos and M. Richardson. Mining the network value of customers. In **KDD’01**, pages 57–66, 2001.
- D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In **KDD’03**, pages 137–146, 2003.
- J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance. Cost-effective outbreak detection in networks. In **KDD’07**, pages 420–429, 2007.
- W. Chen, Y. Wang, and S. Yang. Efficient influence maximization in social networks. In **KDD'09**, pages 199-207, 2009.
- E. Bakshy, D. Eckles, R. Yan, and I. Rosenn. Social influence in social advertising: evidence from field experiments. In **EC'12**, pages 146-161, 2012.
- A. Goyal, F. Bonchi, and L. V. Lakshmanan. Discovering leaders from community actions. In **CIKM’08**, pages 499–508, 2008.
- N. Agarwal, H. Liu, L. Tang, and P. S. Yu. Identifying the influential bloggers in a community. In **WSDM’08**, pages 207–217, 2008.

References(cont.)

- E. Bakshy, B. Karrer, and L. A. Adamic. Social influence and the diffusion of user-created content. In **EC ’09**, pages 325–334, New York, NY, USA, 2009. ACM.
- P. Bonacich. Power and centrality: a family of measures. **American Journal of Sociology**, 92:1170–1182, 1987.
- R. B. Cialdini and N. J. Goldstein. Social influence: compliance and conformity. **Annu Rev Psychol**, 55:591–621, 2004.
- D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri. Feedback effects between similarity and social influence in online communities. In **KDD’08**, pages 160–168, 2008.
- P. W. Eastwick and W. L. Gardner. Is it a game? evidence for social influence in the virtual world. **Social Influence**, 4(1):18–32, 2009.
- S. M. Elias and A. R. Pratkanis. Teaching social influence: Demonstrations and exercises from the discipline of social psychology. **Social Influence**, 1(2):147–162, 2006.
- T. L. Fond and J. Neville. Randomization tests for distinguishing social influence and homophily effects. In **WWW’10**, 2010.
- M. Gomez-Rodriguez, J. Leskovec, and A. Krause. Inferring Networks of Diffusion and Influence. In **KDD’10**, pages 1019–1028, 2010.
- M. E. J. Newman. A measure of betweenness centrality based on random walks. **Social Networks**, 2005.
- D. J. Watts and S. H. Strogatz. Collective dynamics of ’small-world’ networks. **Nature**, pages 440–442, Jun 1998.
- J. Sun, H. Qu, D. Chakrabarti, and C. Faloutsos. Neighborhood formation and anomaly detection in bipartite graphs. In **ICDM’05**, pages 418–425, 2005.

Thank You!

Yang’s Homepage: http://yangy.org

Jie’s Homepage: http://keg.cs.tsinghua.edu.cn/jietang/