Exponentially Twisted Sampling for Structural Network Analysis in Attributed Networks

Providence University, Oct. 15, 2018
Cheng-Shang Chang

Institute of Communications Engineering
National Tsing Hua University

References

- Cheng-Hsun Chang, Cheng-Shang Chang, Chia-Tai Chang, Duan-Shin Lee, and Ping-En Lu, "Exponentially twisted sampling for centrality analysis and community detection in attributed networks," accepted by IEEE Transactions on Network Science and Engineering. Conference version in IEEE ICC, May 2018.
- Cheng-Shang Chang, Duan-Shin Lee, Li-Heng Liou, Sheng-Min Lu, and Mu-Huan Wu, "A probabilistic framework for structural analysis and community detection in directed networks," IEEE/ACM Transactions on Networking, vol. 26, no. 1, pp. 31--46, 2018.

Overview

- Analysis of networked data has received a lot of attention lately from many researchers in various fields, including physicists, mathematicians, computer scientists, biologists, and sociologists.
- In order for researchers from various research disciplines to carry out their scientific research, it is thus essential to have a framework for the analysis of networked data.
- The aim of this talk is to introduce a unified framework for the analysis of networked data, including centrality analysis, community detection and clustering.

Modeling

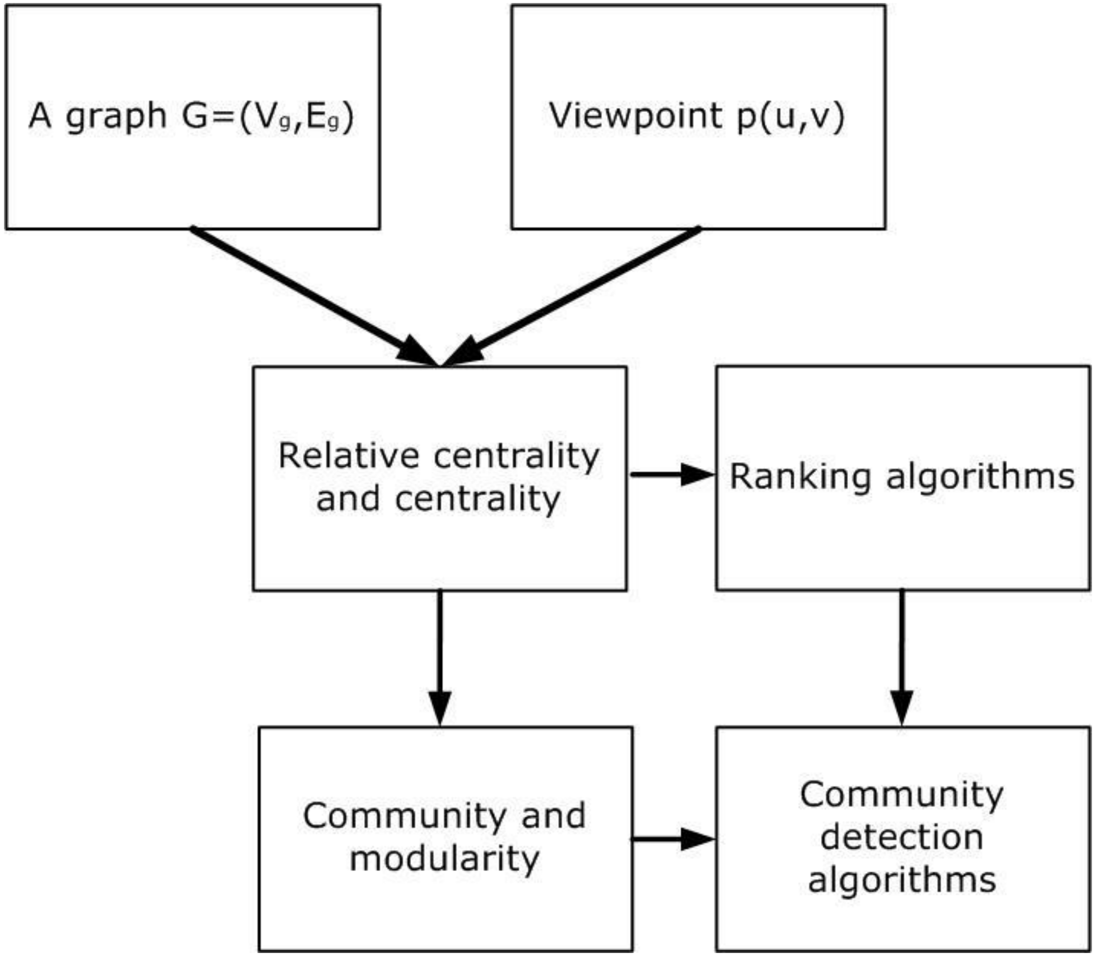
Structural network analysis

- Ranking (centrality): degree centrality, eigenvector centrality, Katz's centrality, PageRank
- Similarity: cosine similarity, Pearson's correlation coefficient, Katz's similarity
- Community detection: Newman's modularity, stability
- Homophily and assortative mixing

Centrality analysis

Community detection

A unified probabilistic framework



Outline

- Introduction
- Probabilistic Framework of Sampled Graphs
  - Sampled Graphs
  - Covariance, Community, and Modularity
- Exponentially Twisted Sampling
- Path Measures in Attributed Networks
  - Influence Centralities
  - Trust Centralities
  - Advertisement-specific Influence Centralities
  - Clustering with a Distance Measure
- Experimental Results
- Conclusion

Outline

- **Introduction**
- Probabilistic Framework of Sampled Graphs
  - Sampled Graphs
  - Covariance, Community, and Modularity
- Exponentially Twisted Sampling
- Path Measures in Attributed Networks
  - Influence Centralities
  - Trust Centralities
  - Advertisement-specific Influence Centralities
  - Clustering with a Distance Measure
- Experimental Results
- Conclusion

Introduction of Attributed Networks (1/2)

- An attributed network is a generalization of a network, which could have node attributes and edge attributes that specify the "features" of nodes and edges.
  - Signed edges: Each edge is labeled with either a positive sign or a negative sign to indicate the friendship/enemy relationship between the two ends of an edge.
  - Every node in the network represents a person with different ratings/interests on various topics.
  - The edge attribute of a directed edge is the distance from one end to the other end of the directed edge.

Introduction of Attributed Networks (2/2)

- An attributed network is a generalization of a graph $G(V, E)$ by assigning each node $u \in V$ an attribute $h_V(u)$ and each edge $e \in E$ an attribute $h_E(e)$.
- As such, an attributed network can be represented as $G(V, E, h_V(\cdot), h_E(\cdot))$, where $h_V(\cdot)$ and $h_E(\cdot)$ are called the node attribute function and the edge attribute function, respectively.

Introduction of Centrality

- Centralities have been widely used for ranking the important nodes in social network analysis.
- In the literature [5], there are various notions of centralities.
  - Degree centrality
  - Eigenvector centrality
  - Katz centrality
  - PageRank
  - Closeness centrality
  - Betweenness centrality

[5] M. Newman, *Networks: an introduction*. OUP Oxford, 2009.

Introduction of Community Detection

- The basic goal of community detection is to find a partition of a graph so as to discover and understand the large-scale structure of a network.

- What a "good community" should look like?

  ILL-POSED PROBLEM !

  - People have different views.
  - Communities can only be formally defined on top of a specific viewpoint.

Framework

Sampled Graph

- Original sampling distribution $p_0(r)$
- Path measure $f(\cdot)$

Exponential Twisting

- Once a specified average path measure $\overline{f}$ (or the inverse temperature vector θ) is given.
- One can have the new sampling distribution $p(\cdot)$ in (18).

Network Analysis

- Bivariate distribution in (7).
- Marginal distributions in (8) correspond to the centrality.
- Community detection algorithms.

Outline

- Introduction
- **Probabilistic Framework of Sampled Graphs**
  - **Sampled Graphs**
  - Covariance, Community, and Modularity
- Exponentially Twisted Sampling
- Path Measures in Attributed Networks
  - Influence Centralities
  - Trust Centralities
  - Advertisement-specific Influence Centralities
  - Clustering with a Distance Measure
- Experimental Results
- Conclusion

Sampled Graph (1/2)

- The main idea in the probabilistic framework for structural analysis in undirected/directed networks [9][23], is to sample a network by randomly selecting a path in the network.
- A network with a path sampling distribution is then called a sampled graph that can in turn be used for structural analysis of the network, including centrality.

[9] C.-S. Chang, C.-J. Chang, W.-T. Hsieh, D.-S. Lee, L.-H. Liou, and W. Liao, "Relative centrality and local community detection," *Network Science*, vol. 3, no. 4, pp. 445–479, 2015.
[23] C.-S. Chang, D.-S. Lee, L.-H. Liou, S.-M. Lu, and M.-H. Wu, "A probabilistic framework for structural analysis and community detection in directed networks," IEEE/ACM Transactions on Networking, 2018.

Sampled Graph (2/2)

- There are many methods for sampling a graph with a randomly selected path.
  - Sampling by uniformly selecting a directed edge.
  - Sampling by a random walk on an undirected network with path length 1 or 2.
  - Sampling by a Markov chain.

Notations for the Sampled Graph (1/2)

- A network is modeled by a graph $G(V, E)$
  - V denotes the set of nodes in the graph.
  - E denotes the set of edges in the graph.
  - $n = |V|$ be the number of vertices in the graph and index the n vertices from $1, 2, \ldots, n$.
  - $m = |E|$ be the total number of edges.
  - k_v be the degree of node v, $v = 1, 2, \ldots, n$.
  - Let $A = (a_{i,j})$ be the $n \times n$ adjacency matrix of the graph, i.e.,

$$a_{i,j} = \begin{cases} 1, & \text{if there is an edge from vertex } i \text{ to vertex } j, \\ 0, & \text{otherwise.} \end{cases}$$

Notations for the Sampled Graph (2/2)

- Let $R_{u,w}$ be the set of (directed) paths from u to w, and $R = \bigcup_{u,w \in V} R_{u,w}$ be the set of paths in the graph $G(V, E)$.
- According to a probability mass function $p(\cdot)$, called the path sampling distribution, a path $r \in R$ is selected at random with probability $p(r)$.

Sampling by Uniformly Selecting a Directed Edge

- Given a directed graph $G = (V, E)$ with the adjacency matrix $A = (a_{i,j})$.
- Sampling by uniformly selecting a directed edge has the following probability mass function:

$$p(r) = \begin{cases} 1/m, & \text{if } r \text{ is an edge from vertex } i \text{ to vertex } j, \\ 0, & \text{otherwise,} \end{cases}$$

Sampling by the Random Walk (1/2)

- A path r with length 1 can be represented by the two nodes $\{u_1, u_2\}$ it traverses. Similarly, a path with length 2 can be represented by the three nodes $\{u_1, u_2, u_3\}$ it traverses.

Sampling by the Random Walk (2/2)

- A random walk with path length not greater than 2 can be generated by the following two steps:
  1. With the probability $\frac{k_v}{2m}$, an initial node v is chosen.
  2. With probability $\beta_i, i = 1, 2$, a walk with length i is chosen, where $\beta_1 + \beta_2 = 1$ and $\beta_i \geq 0$.

  As such, we have

$$p(r) = \begin{cases} \frac{\beta_1}{2m} a_{u_1,u_2}, & \text{if } r = \{u_1, u_2\}, \\ \frac{\beta_2}{2m} \frac{a_{u_1,u_2} a_{u_2,u_3}}{k_{u_2}}, & \text{if } r = \{u_1, u_2, u_3\}, \end{cases} \qquad (1)$$

Bivariate Distribution

- Let U (resp. W) be the starting (resp. ending) node of a randomly select path by using the path sampling distribution $p(\cdot)$. Then the bivariate distribution

$$p_{U,W}(u,w) = \mathsf{P}(U = u, W = w) = \sum_{r \in R_{u,w}} p(r) \qquad (7)$$

is the probability that the ordered pair of two nodes (u, w) is selected.

- As such, $p_{U,W}(u, w)$ can be viewed as a similarity measure from node u to node w and this leads to the definition of a sampled graph.

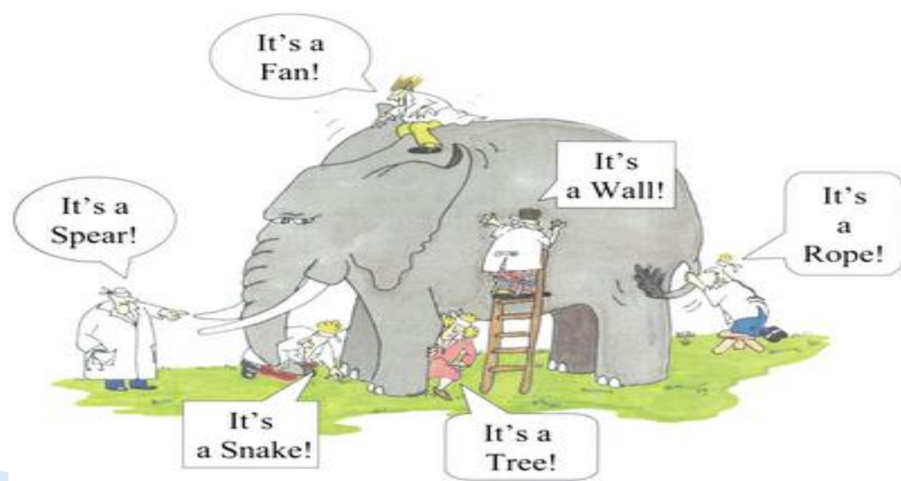
Definition of a Sampled Graph

- A sampled graph is denoted by the two-tuple $(G(V,E), p_{U,W}(\cdot,\cdot))$, which means a graph $G(V,E)$ that is sampled by randomly selecting an ordered pair of two nodes (U,W) according to a specific bivariate distribution $p_{U,W}(\cdot,\cdot)$ in (7).
- Let $p_U(u)$ (resp. $p_W(w)$) be the marginal distribution of the random variable U (resp. W), then $p_U(u)$ can be viewed as a centrality of u.

$$p_U(u) = \mathsf{P}(U = u) = \sum_{w=1}^{n} p_{U,W}(u,w), \tag{8}$$

Sampling a network

- Characterize a graph by a bivariate distribution that specifies the probability of the two vertices appearing at both ends of a "randomly" selected *path in the graph.*
- An old saying, "blind men and an elephant."
- A bunch of blind persons try to figure out what an elephant looks like by touching the elephant with their both hands."



Outline

- Introduction
- Probabilistic Framework of Sampled Graphs
  - Sampled Graphs
  - Covariance, Community, and Modularity
- **Exponentially Twisted Sampling**
- Path Measures in Attributed Networks
  - Influence Centralities
  - Trust Centralities
  - Advertisement-specific Influence Centralities
  - Clustering with a Distance Measure
- Experimental Results
- Conclusion

Exponentially Twisted Sampling (1/10)

- In order to take the attributes of nodes and edges into account, one needs a biased viewpoint.
- As such, the sampling method in an attributed network and the corresponding probability measure needs to be a twisted probability measure of its underlining network.

Exponentially Twisted Sampling (2/10)

- For this, we propose using a path measure $f(r)$ that maps every path r in an attributed network to a real-valued vector.
- By specifying the average values of a path measure, we then have a set of constraints for the twisted sampling probability measure.
- This then leads to the exponentially twisted probability measure that minimizes the Kullback-Leibler (KL) distance between the twisted probability measure and the original probability measure.

Exponentially Twisted Sampling (3/10)

- For a path r that traverses a sequence of nodes $\{u_1, u_2, \ldots, u_{k-1}, u_k\}$ in an attributed network, we can define a path measure $f(r)$ as a function of the attributes of the nodes and edges along the path r, i.e.,

$$\{h_V(u_1), \ldots, h_V(u_k), h_E(u_1, u_2), \ldots, h_E(u_{k-1}, u_k)\}$$

- We assume that a path measure $f(\cdot)$ is a mapping from the set of paths R to a L-dimensional real-valued vector in $\mathcal{R}^L$, i.e.,

$$f(r) = \big(f_1(r), f_2(r), \ldots, f_L(r)\big),$$

where $f_i(\cdot), i = 1, 2, \cdots, L$ are real-valued functions.

Exponentially Twisted Sampling (4/10)

- Now suppose that we have already had a sampled graph $(G(V,E), p_0(\cdot))$ that uses the probability mass function $p_0(r)$ to sample a path r in $G(V,E)$.

- The question is how the sampling distribution should be changed so that the average path measure is equal to a specified vector $\bar{f}$.
- In other words, what is the most likely sampling distribution $p(\cdot)$ that leads to the average path measure $\bar{f}$ given that the original sampling distribution is $p_0(\cdot)$?

Exponentially Twisted Sampling (5/10)

- We introduce the Kullback-Leibler (KL) distance between two probability mass functions $p(\cdot)$ and $p_0(\cdot)$:

$$D(p\|p_0) = \sum_{r \in R} p(r) \log\left(\frac{p(r)}{p_0(r)}\right) \tag{13}$$

- The KL distance is known to be nonnegative and it is zero if and only if $p(\cdot) = p_0(\cdot)$.
- The larger the KL distance is, the more unlikely for $p_0(\cdot)$ to behave like $p(\cdot)$.

Exponentially Twisted Sampling (6/10)

$$\begin{aligned} \min \quad & D(p \| p_0) \\ s.t. \quad & \sum_{r \in R} p(r) = 1, \\ & \sum_{r \in R} f(r) p(r) = \overline{f} \end{aligned} \qquad (14)$$

- The first constraint states that the total probability must be equal to 1.
- The second constraint states that the average path measure must be equal to $\overline{f}$ with the new sampling distribution $p(\cdot)$.

Exponentially Twisted Sampling (8/10)

Thus,

$$p(r) = \exp(\alpha - 1) \times \exp(\theta \cdot f(r)) \times p_0(r) \qquad (17)$$

Since $\sum_{r \in \mathcal{R}} p(r) = 1$, it then follows that

$$p(r) = C \times \exp(\theta \cdot f(r)) \times p_0(r), \qquad (18)$$

where

$$C = \frac{1}{\sum_{r \in \mathcal{R}} \exp(\theta \cdot f(r)) \times p_0(r)} \qquad (19)$$

is the normalization constant.

Exponentially Twisted Sampling (9/10)

- To solve the parameter vector θ, we let
$$F = \log(1/C).$$
- The quantity F is called the free energy as it is analogous to the free energy in statistical mechanics [3]. Also, the parameter vector θ is called the inverse temperature vector.
- It is easy to see that for $i = 1, 2, \cdots, L$ that
$$\frac{\partial F}{\partial \theta_i} = \sum_{r \in \mathcal{R}} f_i(r) p(r) = \overline{f_i} \tag{20}$$

[3] M. Newman, *Networks: an introduction*. OUP Oxford, 2009.

Exponentially Twisted Sampling (10/10)

- To summarize, in order to define the centrality of an attributed network, one needs
  1. The original sampling distribution $p_0(r)$ for the network.
  2. The path measure $f(\cdot)$ of the attributed network.
  3. Once a specified average path measure $\bar{f}$ (or the inverse temperature vector θ) is given, one can have the new sampling distribution $p(\cdot)$ in (18).
  4. This then leads to the bivariate distribution in (7).
  5. The marginal distributions of that bivariate distribution then correspond to the centrality of the attributed network.

Framework

Sampled Graph

- Original sampling distribution $p_0(r)$
- Path measure $f(\cdot)$

Exponential Twisting

- Once a specified average path measure $\overline{f}$ (or the inverse temperature vector θ) is given.
- One can have the new sampling distribution $p(\cdot)$ in (18).

Network Analysis

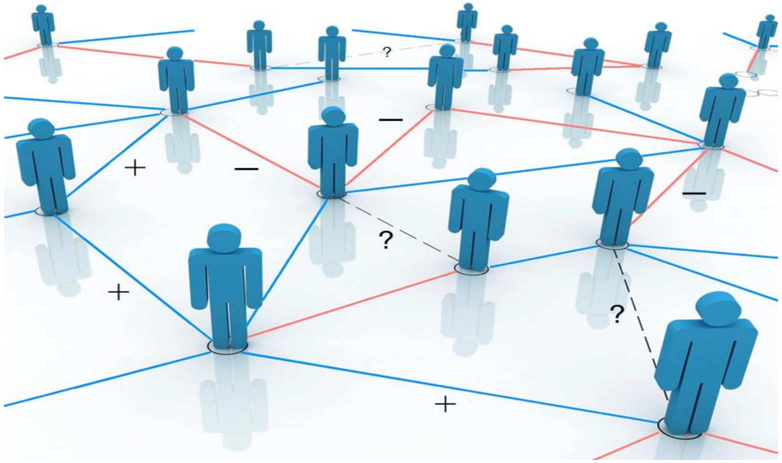
- Bivariate distribution in (7).
- Marginal distributions in (8) correspond to the centrality.
- Community detection algorithms.

Outline

- Introduction
- Probabilistic Framework of Sampled Graphs
  - Sampled Graphs
  - Covariance, Community, and Modularity
- Exponentially Twisted Sampling
- **Path Measures in Attributed Networks**
  - **Influence Centralities**
  - Trust Centralities
  - Advertisement-specific Influence Centralities
  - Clustering with a Distance Measure
- Experimental Results
- Conclusion

Influence centralities (1/9)

- We consider a special class of attributed networks, called signed networks.
- A signed network $G = (V, E, h_E(\cdot))$ is an attributed network with an edge attribute function $h_E(\cdot)$ that maps every undirected edge in E to the two signs $\{+, -\}$.
- The positive (resp. negative) sign is represented by 1 (resp. -1).



Influence centralities (2/9)

- We use opinion dynamics to rank the nodes in signed networks.

- If u and w are connected by a positive (resp. negative) edge, then it is very likely that u will have a positive (resp. negative) influence on w and vice versa.
- We can let the opinion propagate through the entire network and count the (expected) number of nodes that have the same opinion as node u has.
- If such a number is large, then node u has a large positive influence on the other nodes in the network.

Influence centralities (3/9)

- In other words, a node u has a large positive influence if there is a high probability that the other end of a randomly selected path has the same opinion as node u. This then leads us to define the notion of influence centralities for ranking nodes in signed networks.

- The idea is based on the belief that "a friend of my friend is likely to be my friend" and "an enemy of my enemy can be my friend" in [5].

Influence centralities (4/9)

- As such, for a path r that traverses a sequence of nodes $\{u_1, u_2, \dots, u_{k-1}, u_k\}$ in a signed network, we define the following path measure as the product of the edge signs along the path, i.e.,

$$f(r) = \prod_{(u_i, u_{i+1}) \in r} h_E(u_i, u_{i+1}) \tag{21}$$

- Note that $f(r)$ is 1 or -1 as the edge attribute function $h_E(\cdot)$ that maps every undirected edge in E to $\{1, -1\}$.

Influence centralities (7/9)

- If we only select paths with length 1, i.e., $\beta_2 = 0$, then there is a closed-form expression for the influence centrality.

$$C = \frac{m}{m^+e^{\theta} + m^-e^{-\theta}} \tag{24}$$

$$\begin{aligned} p_U(u) &= \sum_{w \in V} p_{U,W}(u,w) \\ &= \frac{(k_u^+ e^{\theta} + k_u^- e^{-\theta})}{2(m^+e^{\theta} + m^-e^{-\theta})}, \end{aligned} \tag{25}$$

  - m^+ (resp. m^-): The total number of positive (resp. negative) edges in the graph.
  - k_u^+ (resp. k_u^-): The number of positive (resp. negative) edges of node u.

Influence centralities (8/9)

- Suppose we require the average path measure $\overline{f}$ to be equal to some fixed constant $-1 < \gamma < 1$.
- From (20) this then leads to

$$\gamma = \overline{f} = \frac{\partial F}{\partial \theta} = \frac{m^{+}\exp(\theta) - m^{-}\exp(-\theta)}{m^{+}\exp(\theta) + m^{-}\exp(-\theta)} \qquad (26)$$

$$\Rightarrow \theta = \ln\left(\sqrt{\frac{m^{-}(1+\gamma)}{m^{+}(1-\gamma)}}\right) \qquad (27)$$

Influence centralities (9/9)

- When $\gamma \to 1$, we have $\theta \to \infty$, and $P_U(u) \to \frac{k_u^+}{2m^+}$. This corresponds to positive degree ranking.
- When $\gamma \to -1$, we have $\theta \to -\infty$, and $P_U(u) \to \frac{k_u^-}{2m^-}$. This corresponds to negative degree ranking.
- When $\gamma = \frac{m^+ - m^-}{m^+ + m^-}$, we have $\theta \to 0$, and $P_U(u) \to \frac{k_u}{2m}$. This corresponds to total degree ranking.

- Thus, different choices of γ lead to different ranking methods.

Outline

- Introduction
- Probabilistic Framework of Sampled Graphs
  - Sampled Graphs
  - Covariance, Community, and Modularity
- Exponentially Twisted Sampling
- **Path Measures in Attributed Networks**
  - Influence Centralities
  - **Trust Centralities**
  - Advertisement-specific Influence Centralities
  - Clustering with a Distance Measure
- Experimental Results
- Conclusion

Trust centralities (1/3)

- Another commonly used interpretation is the trusted/untrusted link.
- "An enemy of my enemy can be my friend" is not suitable for modelling trust .
- A path r that traverses a sequence of nodes $\{u_1, u_2, \ldots, u_{k-1}, u_k\}$ can be trusted if every edge is a trusted link so that there exists a chain of trust.
- In view of this, we consider another path measure as the minimum of the edge signs along the path, i.e.,

$$f(r) = \min_{(u_i, u_{i+1}) \in r} h(u_i, u_{i+1})$$

Outline

- Introduction
- Probabilistic Framework of Sampled Graphs
  - Sampled Graphs
  - Covariance, Community, and Modularity
- Exponentially Twisted Sampling
- **Path Measures in Attributed Networks**
  - Influence Centralities
  - Trust Centralities
  - **Advertisement-specific Influence Centralities**
  - Clustering with a Distance Measure
- Experimental Results
- Conclusion

Advertisement-specific influence centralities (1/5)

- We consider a class of attributed networks that have node attributes. For a graph $G(V,E)$ with the node attribute function $h_V(u)$ that maps every node u to a vector in $\mathcal{R}^L$

$$(h_{V,1}(u), h_{V,2}(u), \cdots, h_{V,L}(u)). \quad (29)$$

NATIONAL TSING HUA UNIVERSITY
Institute of Communications Engineering

Advertisement-specific influence centralities (2/5)

- View the graph $G(V, E)$ as a social network with n users and the attribute vector in (29) as the scores of user u on various topics.
- Suppose an advertisement z can be represented by a vector of scores $(z_1, z_2, \dots, z_L)$ with z_i being the score of the i^{th} topic.
- We would like to find out who is the most influential user in the network to pass on z.

Advertisement-specific influence centralities (3/5)

- As the influence centralities in the previous section, we propose using opinion dynamics through a path.
- To model how much a user "likes" an advertisement, we use the similarity measure from the inner product of the score vector of the user and that of the advertisement.

- For a path r that traverses a sequence of nodes $\{u_1, u_2, \cdots, u_{k-1}, u_k\}$ in the attributed network, we define the following path measure

$$f(r) = \min_{u \in r} [\sum_{i=1}^{L} z_i \cdot h_{V,i}(u)]. \quad (24)$$

Outline

- Introduction
- Probabilistic Framework of Sampled Graphs
  - Sampled Graphs
  - Covariance, Community, and Modularity
- Exponentially Twisted Sampling
- Path Measures in Attributed Networks
  - Influence Centralities
  - Trust Centralities
  - Advertisement-specific Influence Centralities
  - Clustering with a Distance Measure
- **Experimental Results**
- Conclusion

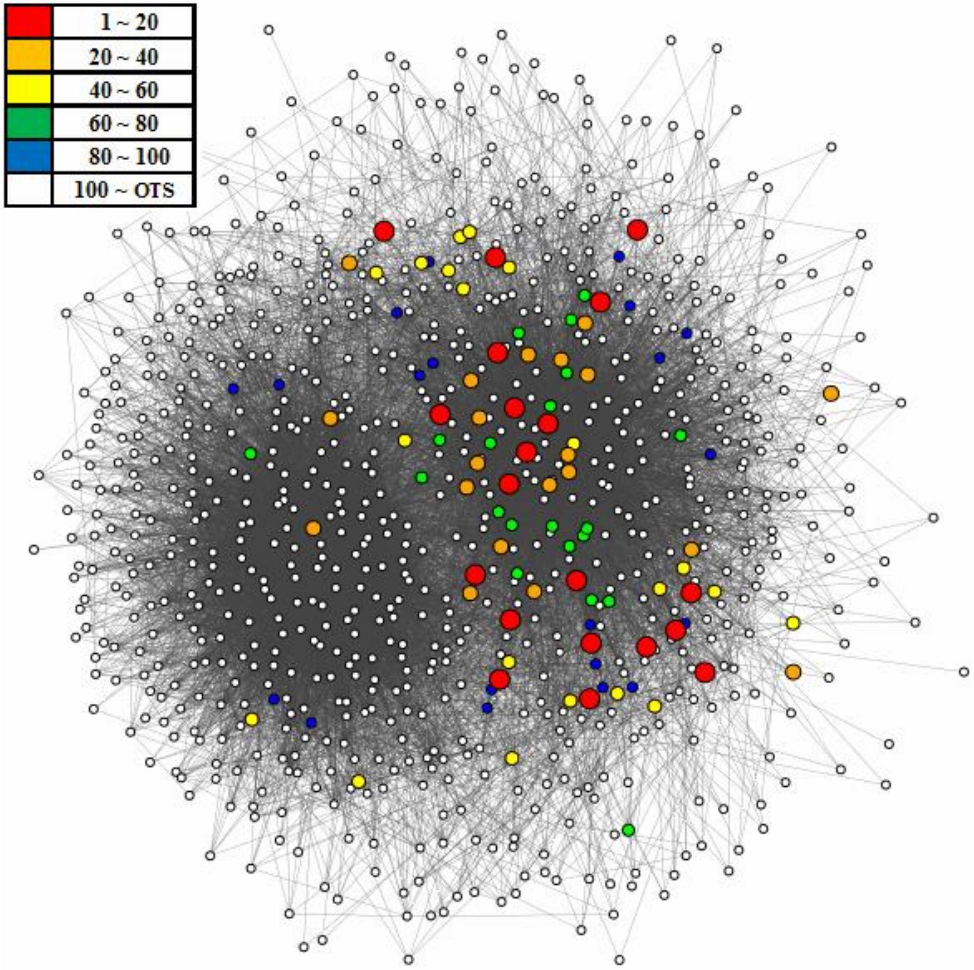
Experimental results for influence centralities in signed networks (1/6)

- Dataset: Political blogs in [27].
  - A directed network of hyperlinks between weblogs collected around the time of the United States presidential election of 2004.
  - 1,490 nodes
  - 19,090 edges
  - These nodes can be partitioned into two (ground-truth) communities (parties).

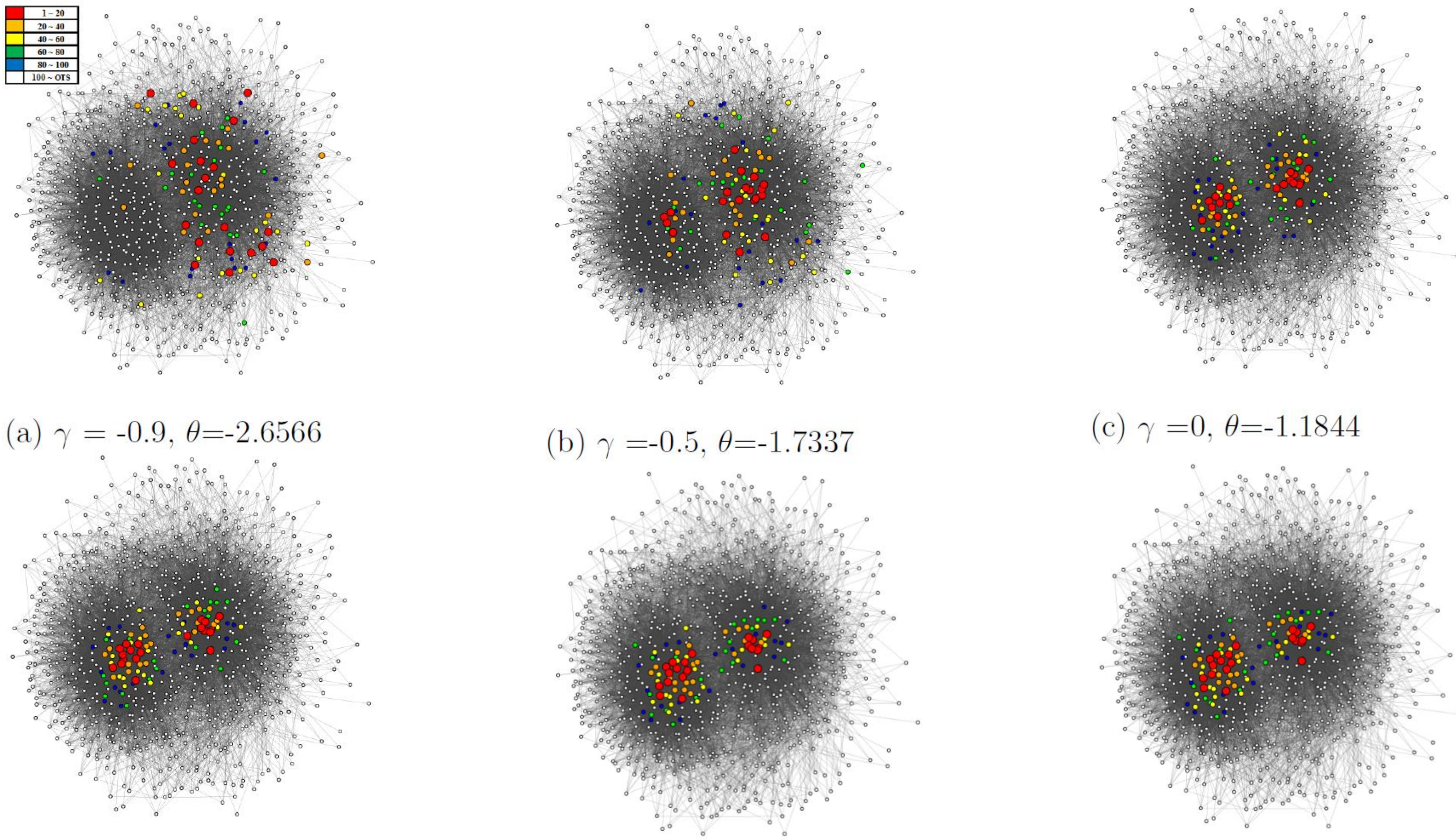
  - Add 4,000 negative edges between two nodes chosen randomly from each community.
  - Symmetrize the adjacency matrix by using the matrix $A + A^T$ to obtain an undirected network.
  - Delete the nodes with degree smaller than 7, remove self edges and multiple edges.
  - Simple undirected signed network with 863 nodes and 16,650 edges, including 15,225 positive edges and 1,425 negative edges.

[27] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 us election: divided they blog," in Proceedings of the 3rd international workshop on Link discovery. ACM, 2005, pp. 36–43.

Experimental results for influence centralities in signed networks (2/6)



Experimental results for influence centralities in signed networks (3/6)



(a) γ = -0.9, θ=-2.6566

(b) γ =-0.5, θ=-1.7337

(c) γ =0, θ=-1.1844

(d) γ =0.5, θ=-0.6351

(e) γ =0.9, θ=0.2878

(f) γ =0.99, θ=1.4623

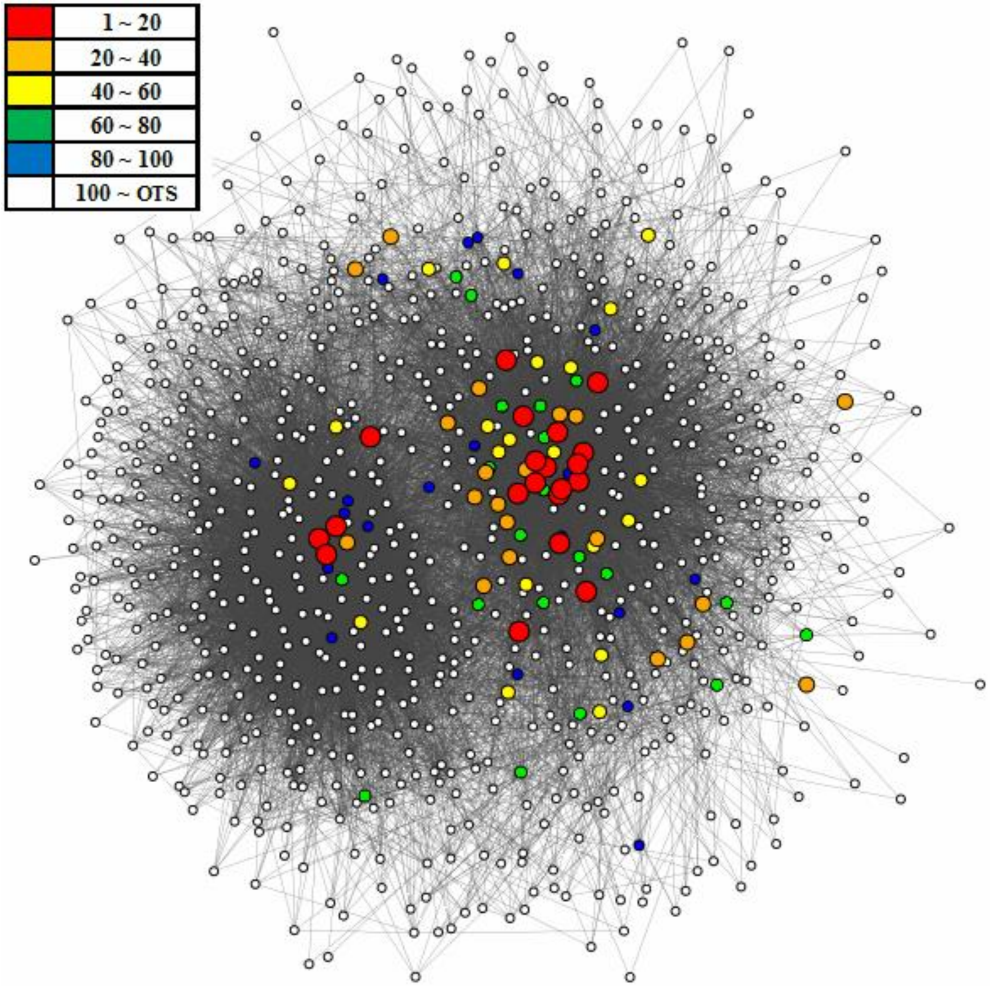
Experimental results for influence centralities in signed networks (4/6)

- The choice of γ is closely related to the three degree ranking methods:
  - Ranking by the number of positive edges (positive degree)
  - Ranking by the number of negative edges (negative degree)
  - Ranking by the total number of edges (total degree).

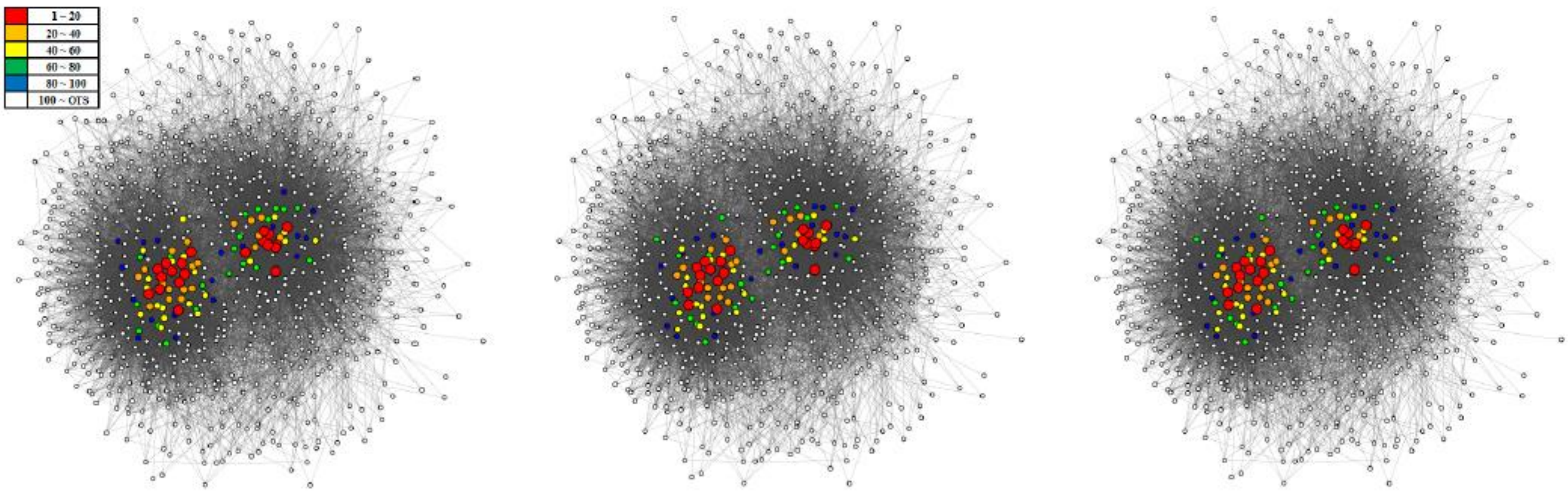
- It is interesting to see that the top 100 nodes gradually move toward the "center" of each community when the value of γ is increased.

Experimental results for influence centralities in signed networks (6/6)

$\beta_1 = 0.7$
$\beta_2 = 0.3$



Experimental results for trust centralities (2/2)



(a) θ =0 (b) θ =1 (c) θ =2

Figure 4.5: The ranking result of trust centralities by β_1=0.7 and β_2=0.3 with different theta

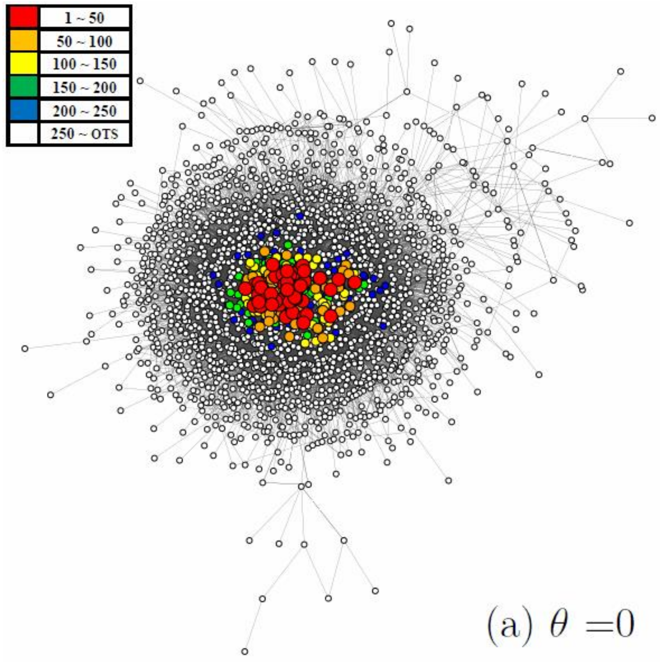
Experimental results for advertisement-specific influence centrality (1/3)

- Dataset: MemeTracker dataset [28]
  - Such a dataset collects the quotes and phrases, called "memes," that frequently appear over time across mass media and personal blogs.
  - We use Carrot2 [40] to classify memes into fifteen topics including "People", "Going", "Know", "Years", "Way", "United States", "States", "Life", "Believe", "Lot", "Love", "America", "Country", "Barack Obama" and "Obama".
  - As "United States" and "Barack Obama" are clearly subsets of the two topics, "States" and "Obama", they are merged into these two topics.
  - Delete the nodes with degree smaller than 7, and remove self edges and multiple edges.
  - 13 topics, 2,082 nodes and 16,503 edges.

[28] J. Leskovec, L. Backstrom, and J. Kleinberg, "Memetracker data," 2008.
[40] S. Osi´nski and D. Weiss, "Carrot2: Design of a flexible and efficient web information retrieval framework," in International atlantic web intelligence conference. Springer, 2005, pp. 439–444.

Experimental results for advertisement-specific influence centrality (2/3)

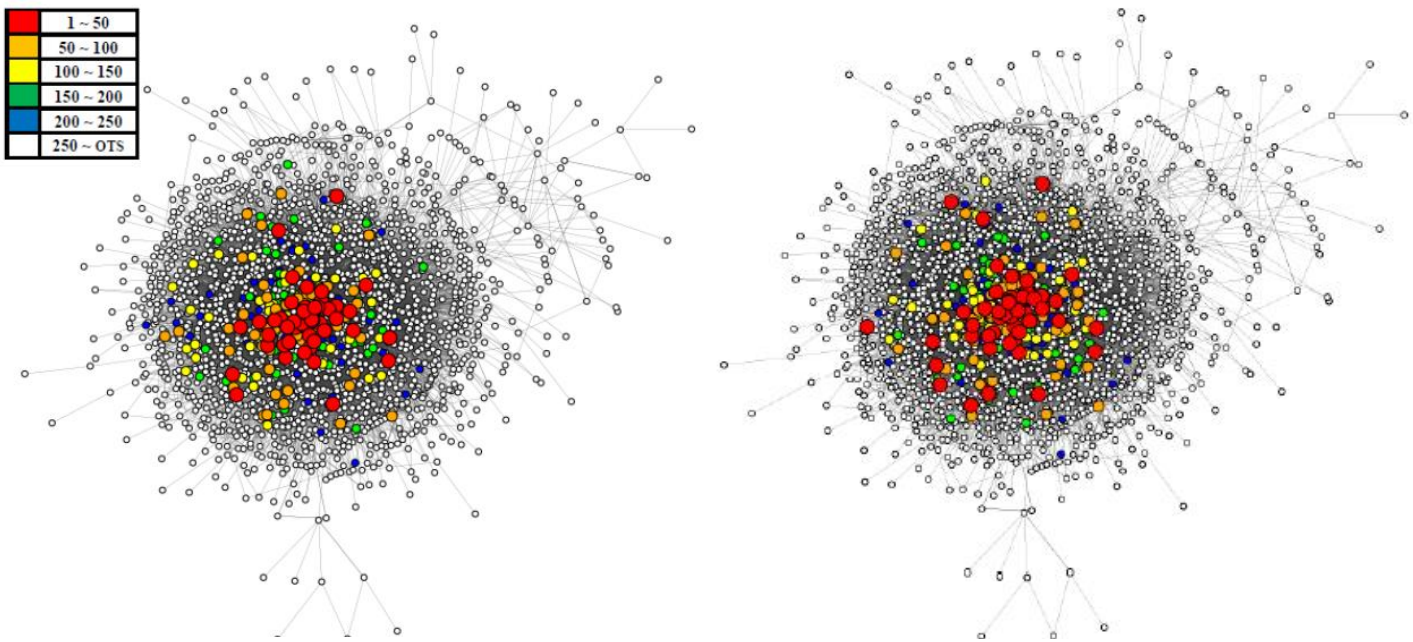
- We choose the phrase "Going" as an advertisement and input it into the existing attributed network that has node attribute.
- We simply use the rank of sampling with path length 1 to do performance evaluation.



(a) θ =0

Experimental results for advertisement-specific influence centrality (3/3)

- Besides, we combine advertisements from different aspects and do the performance evaluation to find the top rank node of these advertisements.



Going and Obama   Going

Conclusion

- We proposed using the exponentially twisted sampling along with path measures for centrality analysis and community detection in attributed networks.

- Provide a general framework for centrality analysis and community detection in attributed networks.