

Using Community Detection to Discover Opinion Leaders in Social Circles

Huajiang Men^a, Xiaoyu Ji^{b,*}, and Wei Wang^b

^a*School of Electronic and Information Engineering, Beihang University, Beijing, 100083, China*

^b*Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing Jiaotong University, Beijing, 100044, China*

Abstract

Discovering opinion leaders in social circles is an important issue in social networks. Most existing opinion leader detection methods usually focus on the whole social network. However, the composition of social networks is complicated, as many social circles or communities based on interests exist in social networks. We find that it is hard to find all the opinion leaders of small social circles if we only focus on the whole network. In this work, we propose a method in which we conduct community detection first and then perform influence analysis on the communities to find the opinion leaders of social circles. Most previous overlapping community detection methods are usually time-consuming and cannot output results in acceptable time on a large-scale dataset; therefore, we propose a linear time complexity overlapping community detection method based on topic graph. We calculate degree centrality, betweenness centrality, closeness centrality, and PageRank value of the nodes in each community detected to find opinion leaders. We collect a large-scale dataset from Zhihu and use it to validate our methods. The extensive results demonstrate that our method can produce better results in finding opinion leaders in social circles compared with other methods.

Keywords: opinion leader; topic; influence; social network; community

(Submitted on October 21, 2018; Revised on November 26, 2018; Accepted on December 25, 2018)

© 2019 Totem Publisher, Inc. All rights reserved.

1. Introduction

In the 1940s, Lazarsfeld [1] first put forward the concept of opinion leader, pointing out that opinion leaders are indispensable intermedia for the information propagation in mass media. According to Merton [2], opinion leaders can also be called the “influentials”. In social networks such as Zhihu, Quora, Weibo, and Twitter, users who usually publish widely accepted opinions, produce valuable content, or have important status in certain fields are more likely to be followed. These users usually have significant influence in their fields and thus can be seen as opinion leaders. Users in social networks can be divided into different fields according to their interest labels, which usually form different communities. It is very possible that these communities vary greatly in size, and they usually have different sets of opinion leaders. If we detect opinion leaders of these communities with different sizes using the same metric, it is likely that some of the opinion leaders in smaller communities cannot be detected. For example, in common social networks, the “Physics” community formed by the “Physics” fans is usually much smaller than the “Movie” community formed by the “Movie” fans. If we put the “Physics” fans and the “Movie” fans in the same large-scale network to analyze and use the same metric to measure their influence, the influence scores of the “Physics” opinion leaders tend to be much lower than those of the “Movie” opinion leaders. Therefore, some of the “Physics” opinion leaders with relatively low influence score may not be detected. In other words, the opinion leaders found in the whole network tend to concentrate in popular fields. In this work, we propose to divide the social network into communities before detecting opinion leaders.

Community structure is an important attribute of social networks. If in a social network the nodes can be easily grouped into (potentially overlapping) sets of nodes such that each set of nodes is densely connected internally, we say that the social network has community structure [3]. A social circle is defined as a group of people who are socially connected. In this paper, we define a social circle in a social network as a community in which users have common interests and certain

* Corresponding author.

E-mail address: 16120380@bjtu.edu.cn

interactive relationships with each other. Community detection algorithms can be divided into overlapping community detection algorithms and non-overlapping community detection algorithms. As users in social networks usually have multiple interests, overlapping community detection algorithms are more suitable for social network community detection based on interests. Most previous overlapping community detection algorithms usually have high time complexity, and the communities detected by these algorithms are not related to user interests. To gain community detection results that are valid for our opinion leader detection method in an accepted time on our dataset, we propose a new community detection algorithm based on topic graph.

We make the following contributions:

- We propose an overlapping community detection algorithm based on topic graph, which has linear time complexity.
- We propose a method of discovering opinion leaders with community detection, which can help in detecting opinion leaders in small social circles.

2. Related Work

Detecting communities in networks is a fundamental issue in network science. In terms of overlapping community detection, Palla et al. [4] proposed the Clique Percolation Method (CPM). In their method, they assume that a community consists of overlapping sets of fully connected subgraphs and detect community structure in the social network by searching for adjacent cliques. There are some other algorithms also based on Clique, such as CPMw [5], SCP [6], and EAGLE [7]. Ahn et al. [8] proposed to use hierarchical clustering of edge similarity to partition links. First, they built a line graph and calculated the similarity between edges. Then, they built a link dendrogram using single-linkage hierarchical clustering and cut this dendrogram at some threshold to get link communities. The time complexity is $O(nk_{max}^2)$, where k_{max} is the maximum degree in the network. To leverage the Label Propagation Algorithm (LPA) in detecting overlapping communities, Gregory et al. [9] proposed the COPRA in which each node has multiple labels. At every step of this method, each node updates its belonging coefficient by averaging the coefficients from all its neighbors. The time complexity is $O(vm \log(vm/n))$ per iteration. Baumes et al. [10] proposed a two-step process. First, the nodes are ranked according to some criterion. Then, highly ranked nodes are iteratively removed until small, disjoint cluster cores formed. Using these cluster cores as seed communities, the second process is to perform an iterative scan and add or remove nodes to expand these cores until the local density function cannot be improved. The worst time complexity of this process is $O(n^2)$. On the foundation of this algorithm, researchers have proposed some methods such as CIS [11], LFM [12], and MONC [13] by modifying the density function.

In previous research, many node influence metrics were put forward. Bonacich et al. [14] firstly proposed the concept of degree centrality, which measures the influence of a node with its degree. This method is simple, intuitive, and has very low time complexity. Freeman et al. [15] proposed closeness centrality. In a graph representing a social network, closeness centrality measures how close a node is to all other nodes in the graph. This method takes the global structure of the network into consideration, and its base stone is the theorem that information propagates fastest along the shortest path. Freeman et al. [16] also proposed betweenness centrality. The betweenness centrality of each node is the number of shortest paths that pass through the node. Bonacich et al. [17] proposed eigenvector centrality (EVC). In this metric, relative scores are assigned to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. Known as a famous web page ranking algorithm, PageRank [18] is also used in social network influence analysis, and it can be seen as a variation of eigenvector centrality. According to Muhammad et al. [19], EVC has the tendency of identifying a set of influential nodes that are all within the same region of a graph. To overcome this shortcoming and improve EVC, they proposed principal component centrality (PCC), in which they use the top n largest eigenvalues and their eigenvectors to compute the centrality of nodes, and this method can output a better result than EVC. Li et al. [20] used an improved closeness centrality to find opinion leaders. To distinguish between positive and negative edges, Robinaugh et al. [21] developed two indices of a node's expected influence (EI) that account for the presence of negative edges. Srinivas et al. [22] proposed Enhanced Degree Centrality. Chen et al. [23] proposed the TopicSimilarRank algorithm based on interactive information and similarity of topics, which can be considered as a variation of PageRank algorithm. Putzke et al. [24] proposed a proxy variable for a node's betweenness centrality that can be calculated in large, evolving networks.

3. Community Detection

In social networks like Zhihu, Quora, and Pinterest, user activities are based on topics. Among the topics, there exists a hierarchical structure, i.e., one topic can be classified into another. It can be seen as a child topic of the topic it belongs to;

and thus a topic may have child topics and one or more parent topics, and naturally it may have descendant topics (including its child topics) and ancestor topics (including its parent topics). For example, the “Machine Learning” topic is a child of the “Artificial Intelligence” topic as well as the “Data Science” topic and the “Computer Science” topic. The topological structure of topics generally forms a directed acyclic graph.

Topics are connected with each other via belonging relationships, constituting the topic graph $T = (V_T, E_T)$, where V_T is the set of topic nodes and E_T is the set of directed edges representing including relationships, $|V_T| = n_T$, $|E_T| = e_T$. Users are connected with each other via follow relationships, constituting the social network $S = (V_S, E_S)$, where V_S is the set of user nodes and E_S is the set of directed edges representing follow relationships, $|V_S| = n_S$, $|E_S| = e_S$. At the same time, there exist follow relationships connecting users and topics. Using these relationships, we can combine topic graph T with the social network S to constitute a user-topic heterogeneous network.

In the hierarchical structure of the topic graph, the lower the level, the fewer the topic nodes at that level. A root node is placed at the lowest level of the graph representing the root topic. The root topic has no practical meaning, but the topic graph is expanded from it.

Our community detection algorithm consists of the following three steps:

3.1. Step 1: Aggregating the Followers

Users in social graph are mapped to the topic graph through follow relationships; therefore, each topic node in the graph possesses a number of followers. Considering the fact that the topics in the topic graph except the root topic are subsets of their ancestor topics, the followers of a topic node are also the followers of its ancestor topics. Therefore, in the first step, we perform depth-first traversal to the topic graph from the root and merge the followers of topic nodes to their ancestors to make topic nodes of the topic graph possess all the followers of their descendants. At the end of this step, all the topic nodes possess not only their own followers but also the followers of their descendant nodes.

3.2. Step 2: Screening the Topics

In the second step, we can divide the whole graph using topics in a certain level of the topic graph. There are very few topics at the first few levels of the topic graph, and the number of topics increases as the level increases. We can get different division granularity by selecting different levels.

It is very possible that topics at the same level overlap with each other because they may have a number of descendants in common. Thus, we need to analyze the overlapping part of the topics and delete the topics that overlap with too many topics in the same level to make the division reasonable. We construct another graph with topics at the level we select, we consider that one topic is subordinated to another topic if the latter contains more than 80% of followers of the former, and then there is an edge from the latter to the former. After finishing the graph construction, we delete the topic nodes in descending order of the nodes' out-degree until only nodes with fewer out-degrees are left. If two topics are subordinated to each other, i.e., the overlap of their follower sets accounts for 80% of the followers of both topics, we believe they are the same topic and only keep one of them.

3.3. Step 3: Finding Largest Weakly Connected Subgraphs

After the above steps, we obtain a few topics and their followers. Then, for all the remaining topics and their followers, we find largest weakly connected subgraphs formed by followers of each topic node and only retain users in these largest weakly connected subgraphs. This step aims to guarantee the connectivity of each community. Then, we will get a division with less overlap.

3.4. Time Complexity Analysis

In the first step of our method, we just perform depth-first traversal to the topic graph. We use an adjacency list, so the time complexity is $O(n_T + e_T)$. The second step is to construct another graph and delete nodes from it. This graph is formed by the topics we want to use to divide the social network, and its size is much smaller, so the time complexity of this step is negligible. In the third step, we traverse the subgraphs formed by the followers of each topic node and find their largest weakly connected subgraphs. The union of all subgraphs that we traverse in this step is roughly equivalent to the whole social network, so the time complexity here is roughly equivalent to traversing the whole social network, which is

$O(n_S + e_S)$. In sum, the aggregate time complexity of our method is $O(n_T + e_T + n_S + e_S)$.

4. Experiment Result

4.1. Community Detection

The dataset used in our experiment is crawled from Zhihu, including 385841 users, 49382 topics, and relationships between users and users, users and topics, and topics and topics.

Zhihu is a question-and-answer social network, where users can share their knowledge, experience, and insights. In September 2017, the number of users of Zhihu exceeded 100 million. All the contents on Zhihu are created by users, including questions, columns, answers, and articles, and all these contents have one or more topic tags. Users in Zhihu choose topics to follow, and Zhihu contents are pushed to users based on the topics they choose. Zhihu topics are managed by users themselves; they can add, delete, merge, or modify topics. Topics are organized in a hierarchical structure, forming a directed acyclic graph with multiple levels. The first level of the topic graph is the root topic. Under the root topic, there are six sub-topics including “Uncategorized Topics”, “Disciplines”, “Entities”, “Metaphysical Topics”, “Industry”, and “Life, Art, Culture, and Activities”, and there are more subdivisions below them. All the topics followed by users are included in this topic graph.

In our experiments, we divide the social network into communities based on the third level of the topic graph. Before constructing the topic-user heterogeneous network, the user-topic relationships are filtered according to users' contribution to the topic. Only user-topic relationships with the contribution degree larger than or equal to 5 are retained; therefore, a number of inactive users are filtered out. Then, we screen the topics to reduce overlap, and as a result, only ten topics including “Politics and Law”, “Professional Development”, “Religion”, “Third Industry (Service Industry)”, “Cultural Industry”, “Art”, “Entertainment”, “Industry”, “Humanity”, and “Technology and Applied Science” are retained. At the end of our algorithm, we find the largest weakly connected subgraphs formed by followers of each of these topics, which constitute communities corresponding to these ten topic categories.

For the dataset of our experiment, the algorithm proposed in this work can output results in less than half an hour. Most other overlapping community detection algorithms, such as CPM, LFM, and EAGLE, cannot output results in an acceptable time. The COPRA algorithm can stop in an acceptable time on our dataset, but its results are not satisfactory: the number of communities is small, and the final results are unstable. More importantly, compared with the results of other community discovering algorithms, the communities discovered in our community detection algorithm have topic labels; thus, our algorithm is more suitable for our opinion leader detection method.

In our experiment, the output of our community detection algorithm is ten communities (as shown in Table 1). Their sizes vary from thousands to tens of thousands.

Table 1. Community detection result

| Topic category | Node number |
|-----------------------------------|-------------|
| Industry | 7678 |
| Third industry (service industry) | 10652 |
| Cultural industry | 9698 |
| Humanity | 15728 |
| Technology and applied science | 29579 |
| Art | 31724 |
| Entertainment | 29710 |
| Professional development | 8068 |
| Religion | 2546 |
| Politics and law | 37159 |

4.2. Influence Analysis

As we can see in Table 1, these communities vary greatly in size. We conduct influence analysis in communities we detected. In our experiment, we calculate in-degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, and PageRank value of the nodes and then compare results of the opinion leader detection after community detection with the results of opinion leader detection in the whole network.

In social networks, it is generally considered that people with many connections are important, which is also the idea of

degree centrality. For social networks based on the follow relationships, the nodes with more in-degrees can be considered more important. Suppose graph $G = (V, E)$ has $n = |V|$ nodes and $m = |E|$ edges. Equation (1) is the formula of in-degree centrality.

$$C_D(v) = \frac{\text{indeg}(v)}{n - 1} \quad (1)$$

We select the calculation results of the smallest community—the “Religion” community, a relatively large community—the “Art” community, and the whole network to perform a comparative analysis.

We divide the in-degree centrality of all nodes into 100 levels and count the number of nodes at each level (as shown in Figure 1). We find that in the distribution of in-degree centrality, most nodes are concentrated at the lowest level, and there is a cliff-like drop between the lowest level and second lowest level. We can see that as the network scale expands, the influence score of nodes becomes more concentrated on the lowest level. We select the top 0.5% of nodes in the ranking of in-degree centrality as opinion leaders and find the opinion leader set of the “Religion” community and the “Art” community, and their intersection with the opinion leader set of the whole network (as shown in Table 2). We find that due to the fact that the size of the “Art” community is relatively large and on the same order of magnitude as the size of the whole social network, all the opinion leaders detected in the “Art” community can be detected in the whole social network. However, only 2/3 of opinion leaders detected in the “Religion” community, which is much smaller than the whole social network, can be detected in the whole social network.

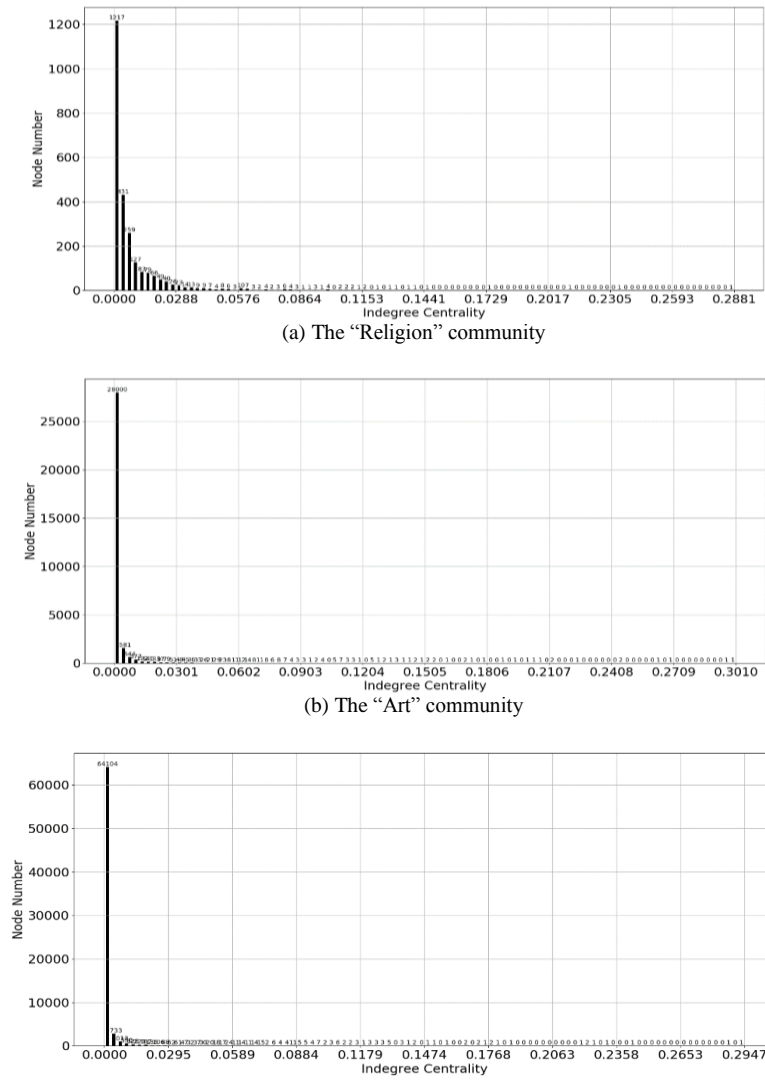


Figure 1. In-degree centrality calculation results

Table 2. Opinion leader detection result using in-degree centrality

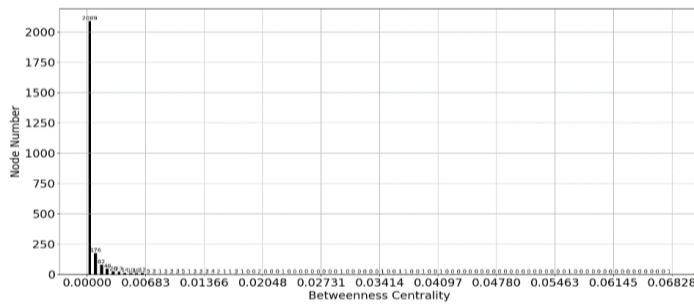
| Community name | Number of opinion leaders | Number of opinion leaders in the whole network | Size of intersection |
|----------------|---------------------------|--|----------------------|
| “Religion” | 12 | 349 | 8 |
| “Art” | 158 | 349 | 158 |

Betweenness centrality is another measure that considers the importance of nodes in connecting other nodes. A way to calculate a node's betweenness centrality is to count the number of shortest paths through the node connecting all pair of nodes, not including the node you are calculating for. Equation (2) is the formula of betweenness centrality.

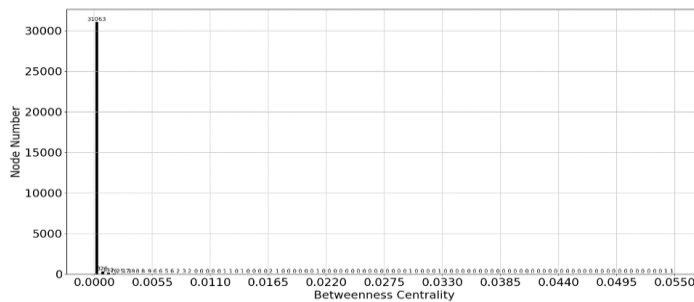
$$C_B(v) = \sum_{s \neq v \neq t} \frac{\delta_{st}(v)}{\delta_{st}} \quad (s, v, t \in V) \quad (2)$$

Where δ_{st} is the number of shortest paths from s to t and $\delta_{st}(v)$ is the number of shortest path through v from s to t .

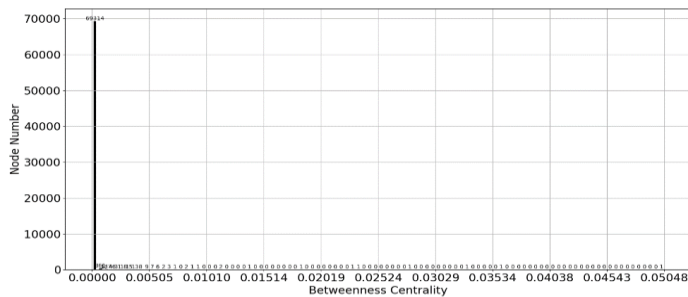
We analyze the betweenness centrality calculation result according to the same procedure as above. We can see in Figure 2 that the betweenness centrality scores of most nodes are also distributed near 0, and nodes with higher scores are very few.



(a) The “Religion” community



(b) The “Art” community



(c) The whole network

Figure 2. Betweenness centrality calculation results

As shown in Table 3, we find that for opinion leader detection in the whole network we achieve a better result using betweenness centrality than the result using in-degree centrality. However, we still cannot find all opinion leaders detected in small communities.

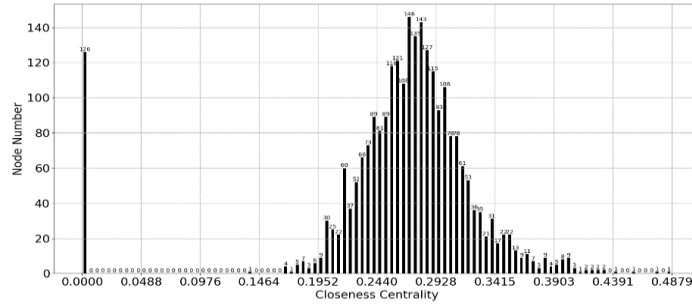
Table 3. Opinion leader detection result using betweenness centrality

| Community name | Number of opinion leaders | Number of opinion leaders in the whole network | Size of intersection |
|----------------|---------------------------|--|----------------------|
| “Religion” | 12 | 349 | 10 |
| “Art” | 158 | 349 | 157 |

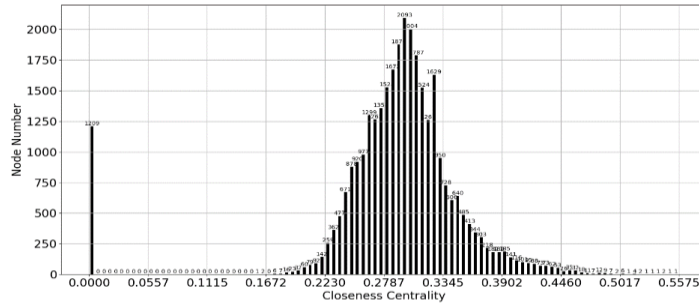
The idea of closeness centrality is that the information of the nodes closer to the center can be transmitted to other nodes faster, and these nodes can be considered more important and thus be seen as opinion leaders. In a more formalized description, influential nodes satisfy the condition that they have a relatively small average length of the shortest paths to other nodes. Equation (3) is the formula of closeness centrality.

$$C_{CL}(v) = \frac{n}{\sum_s d(s, v)}, \quad (s, v \in V) \quad (3)$$

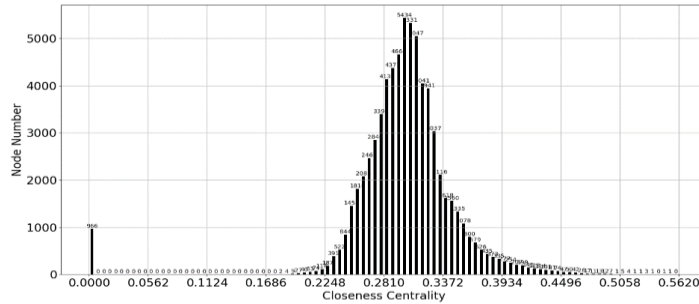
Using the same process to analyze the calculation results, we find that there is a significant difference between the results using closeness centrality and the results using the first two measures. The closeness centrality distribution of nodes is mainly concentrated at the middle levels and the lowest level (as shown in Figure 3). This is contradictory to our intuition that in influence distribution the number of nodes declines with the influence increasing, so we think this measure is not suitable for opinion leader detection.



(a) The “Religion” community



(b) The “Art” community



(c) The whole network

Figure 3. Closeness centrality calculation results

Compared with the idea of degree centrality, which considers nodes with more connections more important, the idea of eigenvector centrality is that nodes with more important neighbors are more important. Equation (4) is the formula of eigenvector centrality.

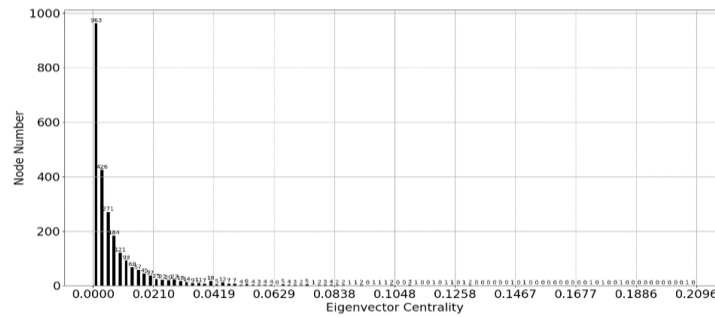
$$C_E(v) = \frac{1}{\lambda} \sum_{t \in M(v)} C_E(t) \quad (4)$$

Where $M(v)$ is the set of node v 's neighbors and λ is a constant. This formula can be rewritten in eigenvector form, which is shown as Equation (5).

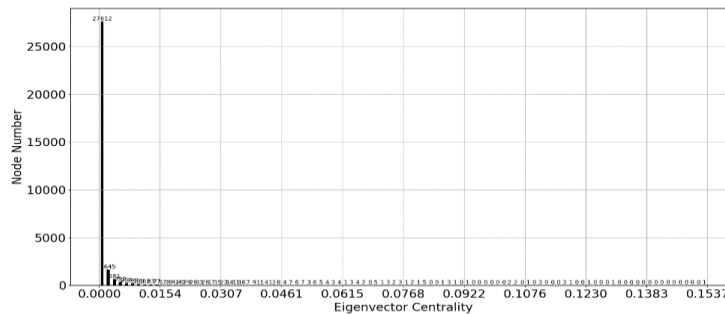
$$Ax = \lambda x \quad (5)$$

Where A is the adjacent matrix of G , λ is the largest eigenvalue of A , and the v th component in vector x is the eigenvector centrality of node v .

Using the same procedure to analyze the calculation result of eigenvector centrality, we find that the analysis result is similar to that of in-degree centrality (as shown in Figure 4 and Table 4).

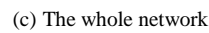
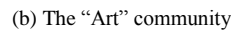
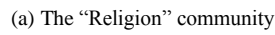


(a) The “Religion” community



| Community name | Number of opinion leaders | Number of opinion leaders in the whole Network | Size of intersection |
|----------------|---------------------------|--|----------------------|
| “Religion” | 12 | 349 | 8 |
| “Art” | 158 | 349 | 158 |

We analyze the calculation result of PageRank values according to the same procedure as above, and the analysis result is similar to that of betweenness centrality (as shown in Figure 5 and Table 5).



| Community name | Number of opinion leaders | Number of opinion leaders in the whole network | Size of intersection |
|----------------|---------------------------|--|----------------------|
| “Religion” | 12 | 349 | 10 |
| “Art” | 158 | 349 | 156 |

To summarize, in our experiment, when we individually discover the opinion leaders in relatively small social circles using the same measures, we can get more opinion leaders of these social circles that meet our criterion, which is a fairer and more accurate result.

5. Conclusions

In this work, we propose a two-step opinion leader discovering method, in which we detect communities first and then conduct influence analysis in each community to discover opinion leaders. Compared with the influence analysis of the whole network, our method can detect opinion leaders more accurately in relatively small communities, in which some opinion leaders may be neglected in the opinion leader detection process of the whole network. For efficiently discovering communities with interest labels, we propose a granularity-controllable community detection method based on topic graph. As an algorithm of linear time complexity, our method significantly reduces the time cost needed for community detection on our large-scale Zhihu dataset.

In the future, we will focus on selecting appropriate topics to detect communities and reducing the overlaps between communities.

Acknowledgements

The work reported in this paper was supported in part by the National Key R&D Program of China (No. 2017YFB0802805), Fundamental Research Funds for the Central Universities of China (No. K17JB00020), and Natural Science Foundation of China (No. U1736114, 61672092).

References

1. P. F. Lazarsfeld, B. Berelson, and H. Gaudet, "The People's Choice," Duell, Sloan & Pearce, Oxford, England, 1944
2. R. K. Merton, "Local and Cosmopolitan Influentials," Perspectives on the American Community, Chicago: Rand McNally, pp. 251-265, 1966
3. Q. Wang and E. Fleury, "Overlapping Community Structure and Modular Overlaps in Complex Networks," Mining Social Networks and Security Informatics, Springer, Dordrecht, pp. 15-40, 2013
4. G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society," *Nature*, Vol. 435, pp. 814-818, June 2005
5. I. Farkas, D. Ábel, G. Palla, and T. Vicsek, "Weighted Network Modules," *New Journal of Physics*, Vol. 9, No. 6, pp. 180, 2007
6. J. M. Kumpula, M. Kivelä, K. Kaski, and J. Saramäki, "Sequential Algorithm for Fast Clique Percolation," *Physical Review E*, Vol. 78, No. 2, August 2008
7. H. Shen, X. Cheng, K. Cai, and M. B. Hu, "Detect Overlapping and Hierarchical Community Structure in Networks," *Physica A: Statistical Mechanics and its Applications*, Vol. 388, No. 8, pp. 1706-1712, April 2009
8. Y. Y. Ahn, J. P. Bagrow, and S. Lehmann, "Link Communities Reveal Multiscale Complexity in Networks," *Nature*, Vol. 466, pp. 761-764, August 2010
9. S. Gregory, "Finding Overlapping Communities in Networks by Label Propagation," *New Journal of Physics*, Vol. 12, No. 10, October 2010
10. J. Baumes, M. K. Goldberg, M. S. Krishnamoorthy, M. Magdon-Ismael, and N. Preston, "Finding Communities by Clustering a Graph into Overlapping Subgraphs," *IADIS AC*, Vol. 5, pp. 97-104, 2005
11. S. Kelley, "The Existence and Discovery of Overlapping Communities in Large-Scale Networks," ProQuest Dissertations Publishing, Ann Arbor, USA, 2009
12. A. Lancichinetti, S. Fortunato, and J. Kertész, "Detecting the Overlapping and Hierarchical Community Structure in Complex Networks," *New Journal of Physics*, Vol. 11, No. 3, March 2009
13. F. Havemann, M. Heinz, and A. Struck, "Identification of Overlapping Communities and their Hierarchy by Locally Calculating Community-Changing Resolution Levels," *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2011, No. 1, January 2011
14. P. Bonacich, "Factoring and Weighting Approaches to Status Scores and Clique Identification," *Journal of Mathematical Sociology*, Vol. 2, No. 1, pp. 113-120, August 1972
15. L. C. Freeman, "Centrality in Social Networks Conceptual Clarification," *Social Networks*, Vol. 1, No. 3, pp. 215-239, 1978
16. L. C. Freeman, "A Set of Measures of Centrality based on Betweenness," *Sociometry*, Vol. 40, No. 1, pp. 35-41, March 1977
17. P. Bonacich and P. Lloyd, "Eigenvector-Like Measures of Centrality for Asymmetric Relations," *Social Networks*, Vol. 23, No. 3, pp. 191-201, July 2001
18. R. Wang, W. Zhang, H. Deng, N. Wang, Q. Miao, and X. Zhao, "Discover Community Leader in Social Network with PageRank," in *Proceedings of International Conference in Swarm Intelligence*, Vol. 7929, pp. 154-162, 2013
19. M. U. Ilyas and H. Radha, "Identifying Influential Nodes in Online Social Networks using Principal Component Centrality," in *Proceedings of 2011 IEEE International Conference on Communications (ICC)*, June 2011
20. L. Yang, Y. Qiao, Z. Liu, J. Ma, and X. Li, "Identifying Opinion Leader Nodes in Online Social Networks with a New Closeness Evaluation Algorithm," *Soft Computing*, Vol. 22, No. 2, pp. 453-464, January 2018
21. D. J. Robinaugh, A. J. Millner, and R. J. McNally, "Identifying Highly Influential Nodes in the Complicated Grief Network," *Journal of Abnormal Psychology*, Vol. 125, No. 6, pp. 747-757, August 2016
22. A. Srinivas and R. L. Velusamy, "Identification of Influential Nodes from Social Networks based on Enhanced Degree Centrality Measure," in *Proceedings of 2015 IEEE International Advance Computing Conference (IACC)*, Vol. 29, No. 2, pp. 1179-1184, July 2015

23. Y. C. Chen, J. Y. Cheng, and H. H. Hsu, "Opinion Leader Mining Algorithm in Microblog Platform based on Topic Similarity," in *Proceedings of 2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, pp. 160-165, Chengdu, 2016
24. J. Putzke and H. Takeda, "Identifying Key Opinion Leaders in Evolving Co-Authorship Networks—A Descriptive Study of A Proxy Variable for Betweenness Centrality," *Complex Networks VII.*, Vol. 644, No. 5, pp. 311-323, March 2016

Huajiang Men is currently a Ph.D. student of engineering in the School of Electronic and Information Engineering at Beihang University. He received his M.S. degree from Dalian Maritime University in 2002. He is mainly engaged in research on mobile communication security such as 3G/4G/5G.

Xiaoyu Ji is currently a M.S. student in the School of Computer and Information Technology at Beijing Jiaotong University. He received his B.S. degree from Shandong Agricultural University in 2016. His main research interests include social networks.

Wei Wang is a full professor in the Beijing Key Laboratory of Security and Privacy in Intelligent Transportation in the School of Computer and Information Technology at Beijing Jiaotong University. He earned his Ph.D. from Xi'an Jiaotong University in 2005. He was a postdoctoral researcher at the University of Trento, Italy, from 2005-2006 and at TELECOM Bretagne and at INRIA, France, from 2007-2008. His research interests include network security and data mining.