

A Signal-Strategy-Based Spectral Clustering Method for Community Detection in Complex Networks

Yutong Cui, Qiang Niu, Zhixiao Wang* and Changjiang Du

School of Computer Science and Technology, China University of Mining and Technology, Xuzhou Jiangsu, 221116, China
zhxwang@cumt.edu.cn

Abstract

The community detection has been one of the core subjects in complex networks. Spectral clustering is an efficient method widely used in this field. In spectral clustering, the Laplacian matrix should be built with similarity matrix, however, similarity matrix is often been replaced by adjacency matrix because few appropriate ways could be used to measure the node similarity in a complex network. As the solution, an appropriate measure of similarity should be proposed to build Laplacian matrix. Signal strategy has been proved to be an efficient method reflecting the relationships between nodes in complex network, and the relationship could be considered as a reasonable scale. This paper presents a semi-supervised spectral approach for community detection, the proposed method uses signal strategy to generate the Laplacian matrix, and utilizes prior knowledge to further guarantee the detection performance. Experiments results showed that the proposed method gave excellent performance on real world network and Lancichinetti-Fortunato-Radicchi (LFR) benchmark, with comparison of other spectral and non-spectral community detection methods.

Keywords: *community detection, spectral clustering, Laplacian matrix, semi-supervised, signal strategy*

1. Introduction

The study of community detection has been one of the core subjects in complex network research and a representative application in data mining. Multiple algorithms has been designed. Girvan and Newman [1] provided a betweenness-based method which divide communities by removing links with the largest betweenness one by one, as GN algorithm. Newman [2] introduced a modularity function Q which measures the quality of a given partition of a network, and improved the GN algorithm into Fast-Newman (FN) algorithm, by searching for the partition with largest Q .

Spectral clustering gains many fundamental advantages comparing to the “traditional algorithms” such as K-means or Single Linkage. By abstracting the structure of the complex network into adjacency matrix, and calculating Laplacian matrix in different model, spectral clustering has proved to be a powerful tool in community detection [3]. White and Smith argued a spectral clustering approach to find communities in graph, Rohe and Chatterjee [4] concentrated on the diversity and superiority of spectral methods in its application in stochastic blockmodel. Mazumder and Tibshirani [5] provided a simple and very efficient semi-supervised spectral clustering method, they extended the spectral regularization to a more sophisticated forms. This paper applies this semi-supervised algorithm to improve the matrix construction, and build Laplacian with guidance of pre-known relationship between nodes.

Received (September 15, 2016), Review Result (January 17, 2017), Accepted (February 6, 2017)

To improve the effect of spectral clustering, there has been many putting their effort on replacing adjacency matrix with other efficient matrix in their research of different fields. Jiang and Jia [6] examined a new Jaccard methods to build similarity matrix. Shi and Fu [7] studied on overlapping community detection based on the similarity of edges. Chen and Wong [8] introduced an algorithm which use topology potential to build Laplacian matrix. As the replacement of adjacency matrix, a suitable similarity-based solution could measure the relationship between nodes more precisely. This provides the inspiration of our algorithm.

Signal Strategy is a well-organized measure which could reveal relationship between nodes as spectral clustering needs. It is first indicated by Hu and Li [9] in field of community identification in complex networks, by combining the strategy with fuzzy C-mean clustering algorithm and F Statistic method. Jiang and Jia [6] also mentioned the signaling method, they used it in NJW algorithm, proved the signal strategy could improve the effect of spectral clustering. Related works have proved that the signal strategy is efficient in revealing neighbor relationship and in helping clustering method end with a reasonable division. The effect of semi-supervised spectral clustering algorithm in community detection could be improved with the Laplacian matrix built with signal matrix.

As summarized from scientists' effort, there exists two restrains in community detection. One is the direct use of adjacency matrix in spectral method, the relationship, or similarity between nodes in specific, might fail to be exposed and end up with an unexpected partition. The other is that prior knowledge might be easy-gaining and useful during the community detection, while traditional spectral method did not make use of them. We consider that the signal strategy and the semi-supervised spectral clustering method could be the solutions correspondingly. Prior labeled knowledge can effectively guide the clustering process, which is the core idea of semi-supervised learning. The signal strategy will result in a good evaluation of similarity.

This paper is organized as follow. In section 2, there will be a detailed description of the algorithm. The experiments and analysis on different benchmark are given in section 3. The conclusion appears in section 4.

2. Method

Given an undirected network $G = \{V, E\}$ consisted of n nodes, $V = \{v_1, \dots, v_n\}$ as the vertex set of G , and E as the edge set. There is adjacency matrix $A = R^{n \times n}$ that the element $A_{ij} = 1$ means the vertex v_i is connected with the vertex v_j and $A_{ij} = 0$ represent no edge between these vertexes. The degree matrix $D \in R^{n \times n}$ is a diagonal matrix in which D_{ii} is the number of neighbor the vertex v_i has, while $D_{ii} = \sum_{k=1}^n A_{ik}$.

The signal strategy could be utilized to build the similarity matrix W together with A , the introduction of signaling is as follow. First a node will be chosen as source randomly and given one unit of signal while the others have no signal. In the next T time steps every node with non-zero signal unit will send its signal to its neighbors and itself. This process will be repeated T times while every sending and receiving are of the same signal quantity the node has, as shown in the Figure 1. An n -dimensional vector of signal quantity record will be established after the process, and we can get n such vectors by choosing every node as the source. Ultimately we standardizing the n vectors, and the distance of each pair will represent the similarity of them. The purpose that we let each node sends signals to itself is to take account of the historical effects, and this has been proved to be helpful to distinguish the amounts of signals between the nodes in the community and outside in a relatively short time period [9]. To maintain the matrix's characteristic of sparseness for spectral clustering, we use the values of signaling only when there is a direct connection between two nodes.

Specifically, if $A_{ij}=1$, we will replace A_{ij} with the distance between v_i and v_j calculated by signaling method, and if $A_{ij}=0$, there won't be any change. We will take the new matrix as the signal similarity matrix $W \in R^{n \times n}$. The degree matrix D will be calculated with W , as $D_{ii} = \sum_{k=1}^n W_{ik}$.

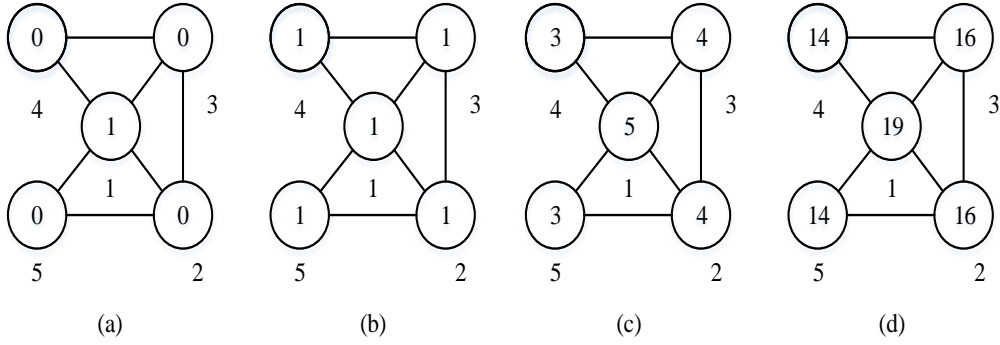


Figure 1. Signal Strategy

(a) We select node 1 as the source and initiate it with 1 signal while others are 0. (b) After the first jump node with signal (node 1) sends its signal to the neighbors. (c) Next, every node with signal will repeat the process with their value remain invariant until the last node finish its work. (d) At the third jump, nodes send the same amount of signals as they received in the last step to their neighbors and themselves.

The form of Laplacian matrix could be various. The standard Laplacian matrix is defined as

$$L = D - W \quad (1)$$

There are two matrices could be called normalized graph Laplacian, both are closely related to each other and are defined as follow [10].

$$L_{sym} = D^{-1/2} L D^{-1/2} \quad (2)$$

$$L_{rw} = D^{-1} L \quad (3)$$

L_{sym} is symmetric matrix while L_{rw} is denoted as it is closely related to random walk. Besides, White and Smith used $W' = D^{-1}W$ as an approximation to L . Chaudhuri *et al.* [11] proposed the regularized graph Laplacian, while Qin and Rohe [12] proposed a regularized spectral clustering (RSC) algorithm based on that. This paper will use L_{rw} as the basic Laplacian structure, as it could improve the result of spectral clustering and is simply for calculation. The final Laplacian is as

$$L_s = D^{-1}(D - W) \quad (4)$$

where W and D are calculated with signal strategy.

The new structure L_s will be utilized to improve the efficiency of semi-supervised spectral clustering which Mavroeidis provided in [4]. This semi-supervised algorithm is as follow. Given a graph with two-cluster structure as V_1 and V_2 , there could be correct label clusters which are the subsets of V_1 or V_2 , as $V_1^{in} \in V_1$ and $V_2^{in} \in V_2$. Define $V^{in} = V_1^{in} \cup V_2^{in}$, and vector t as

$$t = \begin{cases} \sqrt{\frac{d_i}{\text{vol}(V^{in})}}, & \text{if vertex } v_i \in V^{in} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

Where $\text{vol}(V^{in}) = \sum_{k \in V^{in}} d_k$ and d_i is the value of D_{ii} . Define $f(i)$ as

$$f(i) = \begin{cases} \sqrt{\frac{\text{vol}(V_2^{in})}{\text{vol}(V_1^{in})}}, & \text{if vertex } v_i \in V_1^{in}, \\ -\sqrt{\frac{\text{vol}(V_1^{in})}{\text{vol}(V_2^{in})}}, & \text{if vertex } v_i \in V_2^{in}. \end{cases} \quad (6)$$

where $f(i)$ is the addition item which tells the ratio of degree value in different communities. The semi-supervised matrix as follow,

$$L_{semi} = L_s + \gamma t t^T \quad (7)$$

where $\gamma > 0$ is a positive parameter determining the relative importance of the semi-supervised component [3].

After the computation of Laplacian matrix L_{semi} , use classic spectral strategy for community detection. Calculate the eigenvectors $\{x_1, \dots, x_n\}$ of L_{semi} , construct vectors $Y_i (1 \leq i \leq n)$ where Y_i is the i^{th} row of $X = \{x_1, \dots, x_n\}$. Ultimately use the k-means approach to cluster the points $Y_i (i = 1, \dots, n)$.

To serve the conciseness the algorithm will be described as follows.

Algorithm. Semi-supervised spectral clustering with signal strategy

Input:

G : a graph consisting a vertex set V , edge set E and an adjacency matrix A .

Output:

A hard partition of G .

1: Compute the distance between vertex by signaling the network, use the result and A to build the signal similarity matrix W .

2: Compute the signal Laplacian: $L_s = D^{-1}(D - W)$

3: Compute semi-supervised Laplacian: $L_{semi} = L_s + \gamma t t^T$

4: Employ spectral clustering on L_{semi} .

3. Experiments and Results

To investigate the accuracy and efficiency of this algorithm, we choose several spectral and non-spectral methods and test them on both real-world networks and artificial graph. Spectral clustering methods include the basic semi-supervised spectral, standard, SYM, RW and our semi-supervised spectral clustering with signal strategy method. We choose the Fast-Newman (FN) [2] algorithm as a non-spectral comparison, for it is the improvement of the GN[1] method.

The real-world networks in this paper contain Zachary's Karate Club, Dolphin social network, Books about US politics graph and the Power Grid of the Western United States. These datasets could be found on Newman's personal website. Besides, we use

Lancichinetti-Fortunato-Radicchi (LFR) benchmark [13] to create artificial networks and further analyze the quality of our algorithm.

To evaluate the effect, we choose the Modularity Q and NMI as the measurement. Q could be used to measure the quality of a given partition, the larger Q value represents the better partition [2]. Newman defined the modularity as

$$Q = \sum_i (e_{ii} - a_i^2), \quad (8)$$

where e_{ij} means one-half of the fraction of edges in the network that connect vertices in i^{th} group to those in j^{th} group, and a_i is the fraction of all ends of edges that are attached to vertices in i^{th} group.

The normalized mutual information (NMI) index is a measure of similarity between two partitions, which could be used to evaluate the clustering accuracy of a known network. The calculation of NMI is as follow.

$$NMI(A, B) = \frac{-2 \sum_{i=1}^{|A|} \sum_{j=1}^{|B|} N_{ij} \log\left(\frac{N_{ij} N}{N_{i\Box} N_{\Box j}}\right)}{\sum_{i=1}^{|A|} N_{i\Box} \log\left(\frac{N_{i\Box}}{N}\right) + \sum_{j=1}^{|B|} N_{\Box j} \log\left(\frac{N_{\Box j}}{N}\right)}, \quad (9)$$

where A is the standard partition while B representing the experimental one. $|A|$ denotes the number of communities in A , $N_{i\Box}$ the sum of the i^{th} row of N and $N_{\Box j}$ the sum of the j^{th} column of N . All of the experiments were tested by MATLAB.

3.1. Experiments on Zachary's Karate Club

First we consider the well-known karate club friendship network which includes 34 members of a karate club studied by Zachary. Although there is not model answer to its accurate partition, node 1 and 34 are acknowledged as the cores of two different social subnetwork, and we choose these two nodes as the prior knowledge in semi-supervised spectral clustering and our algorithm, with $\gamma=2$ and $k=2$. The main program and parameter initiation are shown below. Other experiments programs of the following subsections are similar.

```
// Read data and build Adjacency Matrix A
A = readGML('karate.gml', 'undirected');
// Hyperparameter initiation
r = 2;
k = 2;
// Prior knowledge
vaArray = [1];
vbArray = [34];
// Calculate Similarity Matrix W with Signal Strategy
W = calSignal(A);
// Calculate Laplacian Matrix L
L = calSimiLaplacian(W, r, vaArray, vbArray);
// Running Spectral Clustering and calculate Q value
result = SpectralClustering(L, k);
Q = calModularityQ(result);
```

We use the modularity Q to evaluate the detecting result for the same realistic scruples, and compare with other spectral methods and Fast-Newman (FN) algorithm. The results are shown in Table 1.

Table 1. Modularity Q for Zachary’s Karate Club

Method	Semi	Semi Signal	Standard	RW	SYM	FN
Q	0.1328	0.3718	0.2340	0.3600	0.3600	0.3718

Notice: RW and SYM represent two different normalized Laplacian as shown in formula (2) (3). FN is the abbreviation of Fast-Newman.

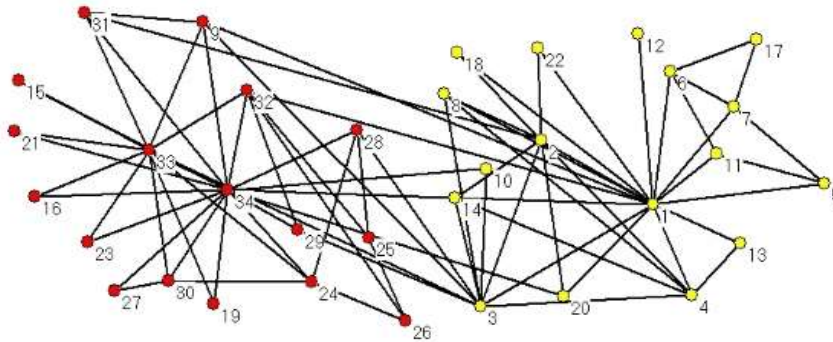


Figure 2. The Clustering Result of Semi Signal Algorithm for Karate Network

The FN algorithm and our approach (Semi Signal) both believed node {3, 10} the members of the group with node 1, while RW and SYM did not. It is clear that our method improve the effect of semi-supervised spectral algorithm under the same setting ($V_1^{in} = \{1\}$, $V_2^{in} = \{34\}$ in both semi-supervised algorithms). Figure 2 is the clustering result of our algorithm.

While running the Semi method on this small scale network, the Laplacian matrix constructed by simple 0-1 adjacency value of nodes may disturb the reflection of relationship between nodes, thus, a weak Laplacian could not end with a satisfied result (larger Q value as we know). However, the signal strategy is good at revealing relationships, and that is the key point of the improvement of Semi Signal.

3.2. Experiments on Dolphin Social Network

In this subsection, we will apply our algorithm on the famous dolphin network, which compile by Lusseau (2003). There are 62 dolphins in this network and can be parted into two groups ($k = 2$). We set $V_1^{in} = \{10\}$, $V_2^{in} = \{31\}$, $\gamma=5$ in both Semi methods. The results are shown in Table 2.

Table 2. Modularity Q for Dolphin Social Network

Method	Semi	Semi Signal	Standard	RW	SYM	FN
Q	0.3848	0.3848	0.3787	0.3787	0.3787	0.3854

The effect of semi method is stable with the increase of network scale. Through there was not improvement to the old Semi method with signal strategy, they both work well and reach a larger Q value than other spectral methods. The FN algorithm reach the largest Q value due to its strategy. The result of our method is shown in Figure 3. The different between semi and non-semi method is the usage of prior knowledge. The results indicates that the well-organized prior information can improve the performance.

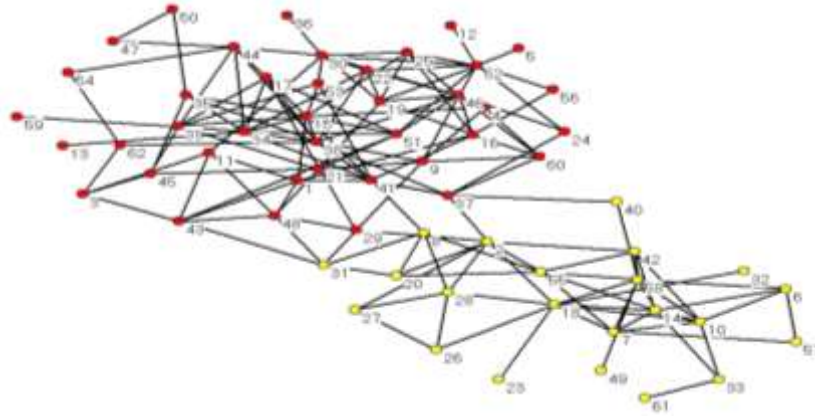


Figure 3. The Clustering Result of Semi Signal Algorithm for Dolphin Network

3.3. Experiments on Books about US Politics Graph

We further tested our method on graph of Books about US politics graph. As the result of dividing the graph into two parts ($k=2$) because spectral clustering has been a hard partition, and set $\gamma=5$. Table 3 show the result of different methods.

Table 3. Modularity Q for Books about US Politics Graph

Method	Semi	Semi Signal	Standard	RW	SYM	FN
Q	0.4282	0.4501	0.4546	0.4546	0.4546	0.4472

It appears that the standard RW and SYM spectral clustering algorithm end with the best Q value, and ours reach the second. The divergence is on node $\{5, 50\}$ and node $\{52\}$, that both algorithms identified that they should belong to different groups, the specific partition plan vary. However, there is an obvious improvement to the traditional Semi function when utilizing the signal strategy. It is clear that Signal Strategy can aid the semi method with the advantage of its reasonable similarity quantification, which could obviously vary the similarity value between different communities. Figure 4 shows the corresponding results.

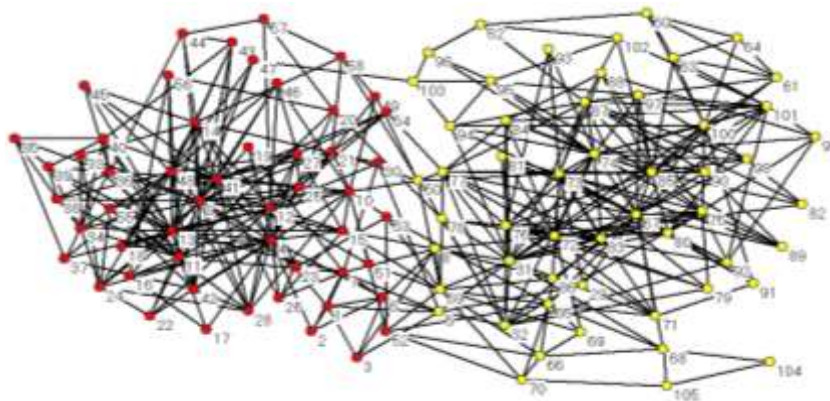


Figure 4. The Clustering Result of Books about US Politics Graph

3.4. Experiments on Power Grid of the Western United States

To analyze the algorithm on network of larger scale, we test it on dataset of power grid of the western United States which compiled by D. Watts and S. Strogatz are available on Newman's website. The Power Grid consists of 4940 nodes. We divided the network into 3

groups ($k=3$) with different algorithm, the effects were evaluated through the calculation of Q value.

Table 4. Modularity Q for Power Grid of Western US

Method	Semi	Semi Signal	Standard	RM	SYM
Q	0.6359	0.6447	0.6322	0.6412	0.6363

By applying different algorithm on Power Grid Network, the Semi Signal method reach the highest Q. We set $V_1^{in} = \{3\}$, $V_2^{in} = \{109\}$, $\gamma=10$ in both Semi methods, and results above indicate the worth of Signal strategy. The FN algorithm does not appear in this table, because its Q value was not within the regular range.

Above comparison on real-world networks shows that once the network's scale become large enough, the prior knowledge could help improve the detection by assisting Semi to reach a higher Q. Signal strategy is effective to evaluate the similarity between nodes and help improve the Semi, as shown in the testing above.

3.5. Experiments on Lancichinetti-Fortunato-Radicchi (LFR) Benchmark

To further evaluate the improvement, we built several artificial networks with LFR. Notice should be taken that when we set a normal average degree, 100 nodes with the average degree of 9 for example, NMI value could meet 1 or near, no matter which algorithm we choose. To distinguish the quality of different methods, we reduce the average degree appropriately and test them with multiple algorithm. First we build two LFR graph with average degree as 3, mixing parameter equal 0.1, one includes 100 nodes while the other 200 nodes. Figure 5 describe the modularity Q and Figure 6 show the NMI value while applying different approach to both subject.

As shown in the results, Semi, Semi Signal and FN algorithm will meet a good partition when the Standard and normalized spectral methods fails. Both Semi functions ends well in LFR100 experiment. The Fast-Newman reach the highest NMI value in LFR200, which means it has the best accuracy of detection, while its Q value is not the largest. Table 5 shows the comparison of these three algorithm.

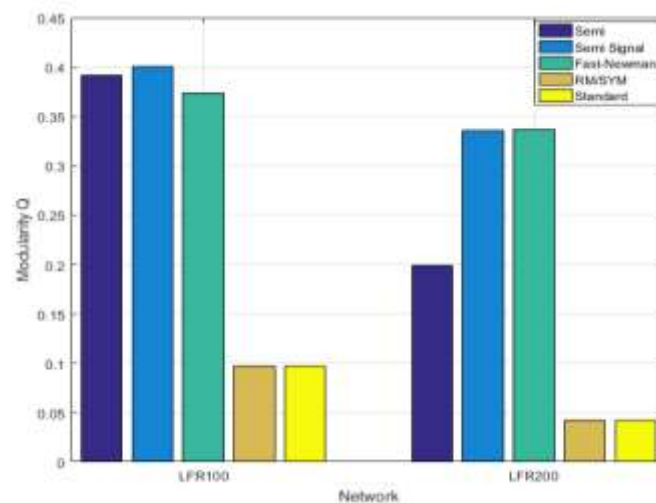


Figure 5. Modularity Q for LFR100 and LFR 200

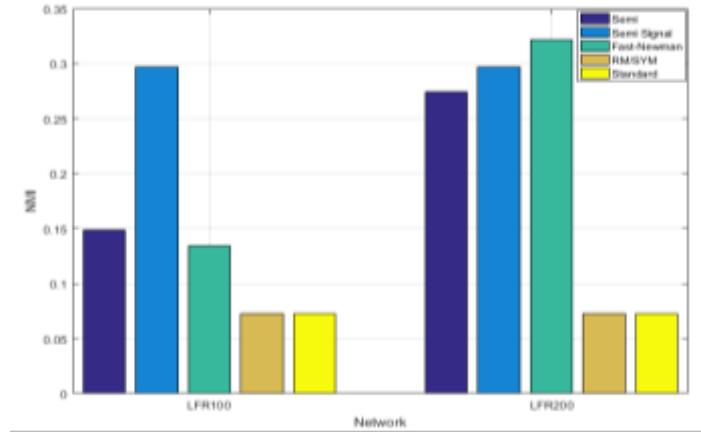


Figure 6. NMI Value for LFR100 and LFR 200

Table 5. Modularity Q and NMI Value for LFR100 and LFR200 Graph

Method	LFR100			LFR200		
	Semi	Semi Signal	FN	Semi	Semi Signal	FN
Q	0.3912	0.4007	0.3737	0.1990	0.3353	0.3364
NMI	0.3484	0.3665	0.1343	0.2741	0.2969	0.3215

It is obvious that signal strategy could improve the effect of semi-supervised spectral clustering, as the old algorithm has advantage in hard situation because it utilize the prior knowledge as V_1^{in} and V_2^{in} . It should be noticed that to simplified the experiments we only tested with single number in each array during the LFR experiments, which means there is only one node in V_1^{in} and V_2^{in} , and we provided the best result under this limitation. Actually there could be more nodes in these array and the results of clustering might be better with different combination, which prove the potential of both Semi algorithms.

To study the influence of parameter γ to the clustering, we tried different γ value from 1 to 10 under the same setting of LFR100 and LFR200 graphs. As shown in Figure 7 and 8, all the results could be stable when γ become large enough, while the signal strategy make the result in LFR200 reach its best with larger γ than Semi method. This might prove that the prior knowledge become more meticulous as the bigger γ value represent a larger proportion the prior knowledge plays, for the result of Semi Signal reach the best with $\gamma = 4$ while the old method with $\gamma = 1$ in LFR200 graph.

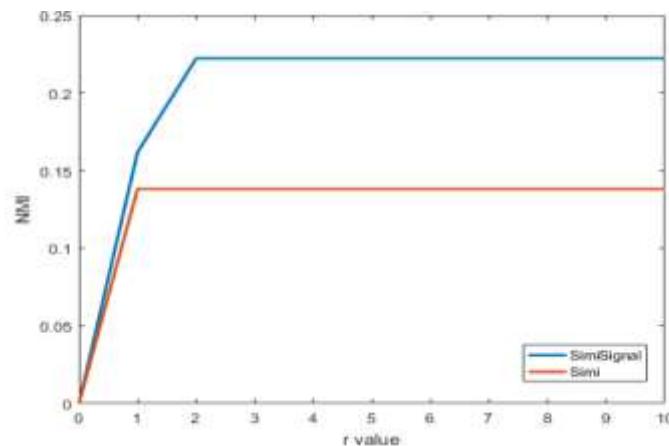


Figure 7. Effect of Clustering Change with γ for LFR100

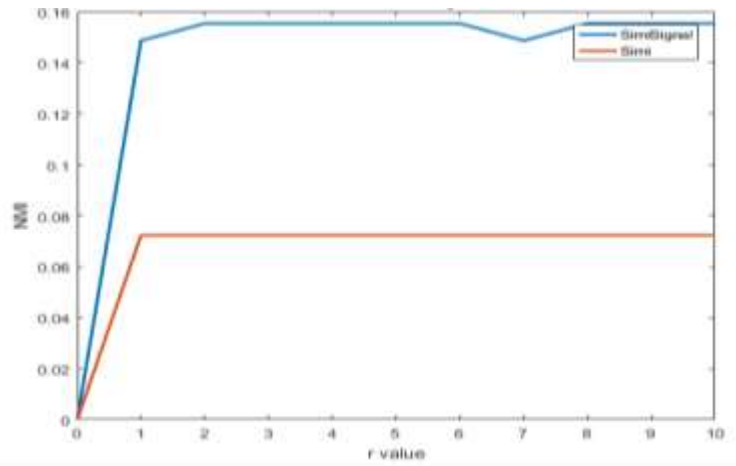


Figure 8. Effect of Clustering Change with γ for LFR200

To further evaluate our algorithm on network with large nodes, we used LFR to build more graph with nodes up to 3000. Figure 9 indicates the improvement of our algorithm. There is no absolute representativeness of every network in the world with such the size though, the fact that signal strategy could improve the effect of semi-supervised spectral clustering is clear.

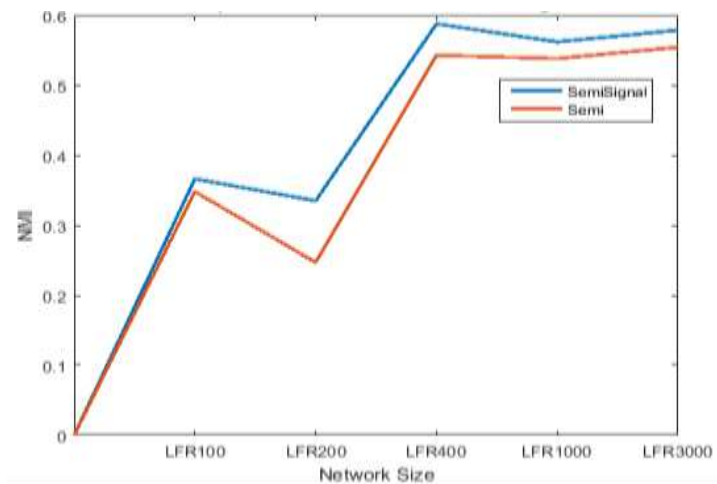


Figure 9. Test on Different Size of Network for Simi and Simi Signal

4. Conclusion

The Semi method has been a prospective clustering approach since inserting the prior knowledge into the detection, which provides a good choice when traditional algorithms work unexpectedly. Our method introduces the successful signal strategy to Semi function, and improves its clustering result while keeping the advantage. By combining the semi-supervised learning theory and signal transforming strategy, the effect of community division in social network has been improved as shown in the experiments.

The insufficiency is that it cost much time while computing the signal matrix. Time complexity could reach $O(MN)$ for a network with N nodes and M edges as signal jumps, although we could reduce the time cost by only calculating the connected nodes from the signal source, it remains the slowest part of the algorithm while facing a large and complex network. Besides, different prior knowledge lead to distinctive result, the condition that user holds might not always provide the best guidance to the algorithm. This could be the common point for any semi-supervised algorithm. The future research will concentrate on

accelerating the algorithm by simplifying the process, dropping the possible redundancy, and finding easier way to determine the prior knowledge for social community detection.

Acknowledgement

This work was supported in part by National Key Research and Development Program (No. 2016YFC060908), the National Natural Science Foundation of China (No.61402482, No.51674255), China Postdoctoral Science Foundation (No.2015T80555), Jiangsu Planned Projects for Postdoctoral Research Funds (No.1501012A), and Qing Lan Project of Jiangsu Province.

Reference

- [1] M. Girvan and M. E. Newman, "Community structure in social and biological networks", Proceedings of the national academy of sciences, vol. 99, no. 12, (2002), pp. 7821-7826.
- [2] M. E. Newman, "Fast algorithm for detecting community structure in networks", Physical review E, vol. 69, no. 6, (2004), pp. 066-133.
- [3] X. Ma and L. Gao, "Non-traditional spectral clustering algorithms for the detection of community structure in complex networks: a comparative analysis", Journal of Statistical Mechanics: Theory and Experiment, vol. 05, (2011), pp. 050-12.
- [4] K. Rohe, S. Chatterjee and B. Yu, "Spectral clustering and the high-dimensional stochastic blockmodel", The Annals of Statistics, (2011), pp. 1878-1915.
- [5] R. Mazumder, T. Hastie and R. Tibshirani, "Spectral regularization algorithms for learning large incomplete matrices", Journal of machine learning research, (2010), pp. 2287-2322.
- [6] Y. Jiang, C. Jia and J. Yu, "Community Detection in Complex Networks Based on Vertex Similarities", Computer Science, vol. 38, no. 7, (2011), pp. 185-189.
- [7] W. Shi, H. Fu and C. Zhang, "Overlapping Community Detection Based on Similarity of Edge", Application Research of Computers, vol. 30, no. 1, (2013), pp. 221-223.
- [8] Z. Wang, Z. Chen, Y. Zhao and S. Chen, "A community detection algorithm based on topology potential and spectral clustering", The Scientific World Journal, (2014), pp. 329-325.
- [9] Y. Hu, M. Li, P. Zhang, Y. Fan and Z. Di, "Community detection by signaling on complex networks", Physical Review E, vol. 78, no. 1, (2008), pp. 016-115.
- [10] L. Huang, R. Li, H. Chen, X. Gu, K. Wen and Y. Li, "Detecting network communities using regularized spectral clustering algorithm", Artificial Intelligence Review, vol. 41, no. 4, (2014), pp. 579-594.
- [11] K. Chaudhuri, F. C. Graham and A. Tsiatas, "Spectral Clustering of Graphs with General Degrees in the Extended Planted Partition Model", In COLT, vol. 23, (2012), pp. 35-1.
- [12] T. Qin and K. Rohe, "Regularized spectral clustering under the degree-corrected stochastic blockmodel", In Advances in Neural Information Processing Systems, (2013), pp. 3120-3128.
- [13] A. Lancichinetti, S. Fortunato and F. Radicchi, "Benchmark graphs for testing community detection algorithms", Physical review E, vol. 78, no. 4, (2008), pp. 046-110.

Authors



Yutong Cui, he was born on 1993 in China. He is a Master of Computer Engineering in Computer Science and Technology. China University of Mining and Technology, Xuzhou, China. His main research interests are social network analysis, machine learning.



Qiang Niu, he was born on 1974 in China. He is also a Ph.D. of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China. He is a Professor and the associate dean with the Department of Computer Science and Technology. He is a Deputy Director of the Mine Information Engineering Research Center in Chinese Coal

Industry. He has published more than 10 academic papers retrieved by SCI. His Main research interests are intelligent information processing, artificial intelligence and pattern recognition, machine learning and data mining and other aspects of research.



Zhixiao Wang, he was born in 1979 in China. He has Ph.D. degree in the Department of Computer Science and Engineering at Tongji University in 2011. He is an Associate professor in the School of Computer Science and Technology, China University of Mining Technology. He has published more than 20 papers in international conferences and journals. His research interests include field theory application, and social network analysis.



Changjiang Du, he was born in 1992 in China. He is a Master of Computer Engineering in School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China. His Main research interests is social network analysis