



Uncovering the overlapping community structure of complex networks by maximal cliques

Junqiu Li, Xingyuan Wang*, Yaozu Cui

Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China

HIGHLIGHTS

- Introduce the maximal cliques, overlapping vertex, bridge vertex and isolated vertex.
- Propose a new algorithm to find all the maximal cliques.
- Expand the maximal cliques by some rules given in this paper.
- Give satisfactory experimental results.

ARTICLE INFO

Article history:

Received 5 March 2014

Received in revised form 9 July 2014

Available online 20 August 2014

Keywords:

Complex networks

Overlapping communities

Maximal cliques

ABSTRACT

In this paper, a unique algorithm is proposed to detect overlapping communities in the un-weighted and weighted networks with considerable accuracy. The maximal cliques, overlapping vertex, bridge vertex and isolated vertex are introduced. First, all the maximal cliques are extracted by the algorithm based on the deep and bread searching. Then two maximal cliques can be merged into a larger sub-graph by some given rules. In addition, the proposed algorithm successfully finds overlapping vertices and bridge vertices between communities. Experimental results using some real-world networks data show that the performance of the proposed algorithm is satisfactory.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

In nature and society, many real networks can be described as complex networks [1–3], e.g. relationship in social system, metabolism in biological systems, protein interaction network, literature reference network, and scientific research cooperation network, etc. Complex networks are a collection of vertices and edges, the vertices of the graph represent the entities of the system and the links represent the interactions between the entities. As the further study on the characters and the physical meanings of the complex networks in the microscopic and mesoscopic level, researchers found that a common feature which is called community structure exists in many real networks. Communities are groups or clusters of vertices within which the network connections are densely connected to each other, but between which they are sparsely connected with the rest of the network [4–7].

In recent years, many algorithms for detecting communities have been proposed. Two classical algorithms are the Kernighan–Lin algorithm that uses a greedy algorithm and classifies the networks by optimizing the number of within and between-community edges [8] and the spectral bisection algorithm based on the eigenvectors of the Laplacian matrix of graph [9]. In Ref. [10], Newman and Girvan proposed a divisive algorithm called the GN algorithm that recursively removes edges with the highest betweenness of the edges, constructed a hierarchical tree, and then we can obtain various divided

* Corresponding author.

E-mail addresses: meiliqitian@126.com (J. Li), wangxy@dlut.edu.cn (X. Wang), cyz3471@sina.com (Y. Cui).

communities by cutting the hierarchical tree at different points. This method makes up for the deficiency of the traditional algorithms, but also has its own defects. This algorithm has high time complexity and has not the defining quantity of community structure in the network. To measure the quality of network partitioning, the modularity [10] represented by the Q function was introduced. Recently, many algorithms are proposed based on modularity. In 2004, Newman proposed a fast greedy algorithm [11] to maximize the modularity, starting with the initial network of n independent communities, with each community having one vertex. Clauset and Newman described a new algorithm for inferring community structure from networks by greedily optimizing the modularity [12], which runs in time $O(md \log n)$ for a network with n vertices and m edges where d is the depth of the dendrogram. This algorithm is considerably faster than most previous traditional algorithms. The same algorithm was presented by Duch and Arenas [13] based on an extremal optimization of the value of modularity. Newman proposed an algorithm using the eigenvectors of matrices [14]. Wang et al. [15] proposed a very fast algorithm only based on local information to detect community. Chen et al. [16] presented a fast and efficient algorithm by mining a node with the closest relations with the community recursively until obtaining a local optimal community.

All the above algorithms mentioned are unable to find overlapping communities, and assume that each vertex belongs to only one community. However, there are some vertices belonging to more than one community in some real networks. So there are a lot of algorithms to explore overlapping community structures in the complex networks recently [17–20]. In addition there is another special type of complex network, which is made up of a lot of maximal cliques. A maximal clique is a complete sub-graph which is not a subset of any other complete sub-graph in a complex network. The number of vertex in a maximal clique is more than two and any two vertices have a link connecting to each other. All the maximal cliques are connected by some overlapping vertices or bridge vertices (in Sections 2.1 and 2.2). This special type of network can be seen in real life. For instance, in scientific collaboration network [21], some authors belong to different research groups characterized by different research directions. Such inter-community authors are often of the most interest. In Chinese medical prescription network (see Fig. 1), a number of Chinese medicinal materials constitute some Chinese medical prescriptions. Fig. 1 is a small medical prescription network, which is an un-weighted network with 20 Chinese medicines as vertices and 6 maximal cliques representing the number of the Chinese medical prescriptions. According to Chinese medical prescriptions, these 6 maximal cliques cannot be merged into a larger sub-graph, so 6 communities and 6 overlapping vertices are obviously detected from Fig. 1.

In this paper, a new algorithm based on the deep and bread searching for extracting all the maximal cliques (complete sub-graphs) is proposed firstly. According to the characteristics of real networks and some given rules, some maximal cliques connected by some overlapping vertices can be merged into a larger sub-graph. All maximal cliques are the community structure in some real networks which are completely composed of the maximal cliques connected by some overlapping vertices. Last, some real world networks are used to test the performance of the presented algorithm.

2. Preliminary concepts

2.1. Overlapping vertex

In the past few years, there are many algorithms to detect non-overlapping communities in complex network. But many real networks have overlapping communities and some vertices belong to more than one community. So this vertex is called the overlapping vertex, and this network having the overlapping vertices is called overlapping neighboring network in this paper. In Fig. 1, this is overlapping neighboring network with six overlapping vertices and six maximal cliques.

2.2. Bridge vertex

Recent studies revealed that some real networks cannot be able to be divided into two communities without allowing bridge vertices connecting multiple groups or communities. The concept of “bridge” is defined as the vertices that cross structural holes between discrete groups of people [22].

To emphasize the importance of bridge vertices in community detection and to illustrate the concept, we take a simple graph shown in Fig. 2 as an example. A visual inspection of this graph most likely suggests three maximal cliques $\{1, 2, 3\}$, $\{4, 5, 6, 7, 8\}$, $\{14, 15, 16, 17\}$. Two maximal cliques are connected by vertex 11, which does not belong to any maximal clique. In this paper, similar vertex 11 is called bridge vertex.

2.3. Isolated vertex

In addition to extracting maximal cliques, overlapping vertex, and bridge vertices, some vertices with low vertex degree do not belong to any of maximal cliques. These vertices are called isolated vertex in this paper. In different real networks, isolated vertices can be isolated from the entire network during the early stage of extracting maximal cliques. At the last stage, isolated vertices can be divided into some communities by some given rules. In Fig. 2, vertex degree of vertices 9, 10, 12 and 13 is 1 and these vertices do not belong to any maximal cliques. So these vertices are called isolated vertices.

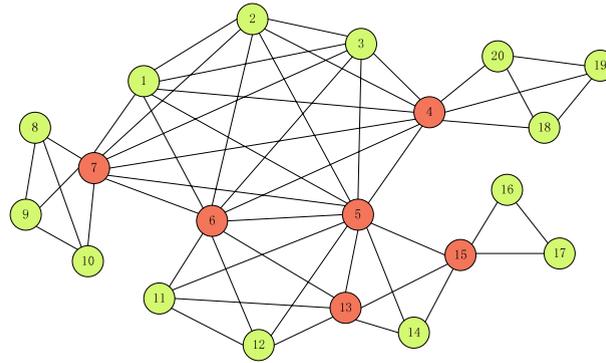


Fig. 1. A small network of Chinese medical prescriptions.

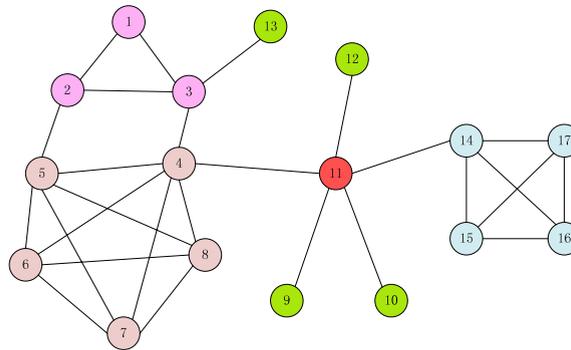


Fig. 2. A small cooperative network with 17 vertices.

3. Algorithm description

In this paper, the algorithm proposed consists of two main procedures. One is extracting all the maximal cliques based on deep and bread searching, and the other is an improving procedure in order to expand the maximal cliques and improve the quality of community structure further.

3.1. Extract all maximal cliques

Scientific collaboration network and Chinese medical prescriptions network have the same network structure, which are composed of many maximal cliques connected by some overlapping vertices. The number of the maximal clique in these networks and the number of each vertex belonging to maximal clique have been obviously important to understand the physical structure of these networks. In scientific cooperation network, each maximal clique represents cooperative outcome or project. The number of maximal cliques shows the number of cooperative outcomes or projects. The number of vertices in each maximal cliques shows the number of scientists who collaborate together. The number of each vertices belonging to a maximal clique explains the number of scientists participating projects. In Chinese medical prescription network, each maximal clique represents a Chinese medicine. The number of vertices in each maximal clique shows the number of the medicinal materials making up a Chinese medicine.

Therefore, detecting communities in these real networks having a certain number of maximal cliques, the vital problem is to accurately extract all maximal cliques from a given network. We can calculate the number of maximal cliques, the number of vertices including in each maximal clique and the number of each vertex belonging to a maximal clique. A unique algorithm based on deep and bread searching is proposed in this paper to extract all maximal cliques from some real networks with a certain number of maximal cliques.

A real network can be modeled as $G(V, E)$, where V represents the set of n vertices or nodes, and E is a set of m links. In this work we shall use un-weighted or weighted complex network. The proposed algorithm based on the deep and bread searching for extracting all the maximal cliques evolves as follows:

- (a) Initially, compute vertex degree of every vertex and remove isolated vertices with vertex degree 1.
- (b) Then, select a vertex $[V_0]$ with the largest vertex degree and find its neighboring vertices.
- (c) Let vertex $[V_0]$ be the starting vertex, searching any neighboring vertex $[N_1]$ with label “ $DV = 1$ ”, then search any neighboring vertex $[N_2]$ of vertex $[N_1]$ and $[V_0]$ with label “ $DV = 1$ ”; search any neighboring vertex $[N_3]$ of vertex $[N_2]$ and $[V_0]$ with label “ $DV = 1$ ” until all neighboring vertices of vertex $[V_0]$ are searched and get back to initial vertex $[V_0]$.

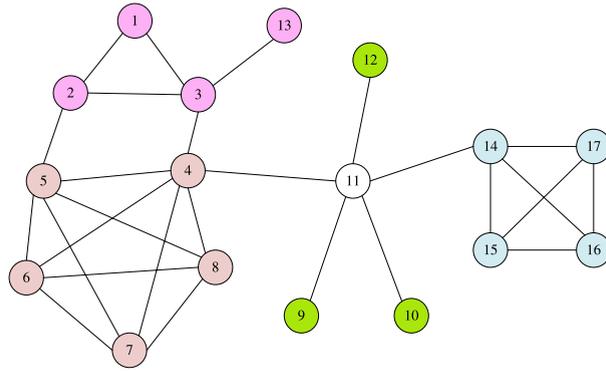


Fig. 3. Three communities in the small cooperative network with 17 vertices.

- (d) After the step (c) get a ring $\{V_0, V_1, V_2, \dots, V_j, V_0\}$, if vertices $\{N_1, N_2, \dots, N_j\}$ are the neighboring vertices of the vertex $[V_0]$, let vertex $[N_1]$ as the starting vertex, searching a part of its neighboring vertex $\{N_3, N_4, \dots, N_j\}$, and let $[N_2]$ be the starting vertex, searching a part of its neighboring vertices $\{N_4, N_5, \dots, N_j\}$. Until vertex $[N_{j-2}]$ is a neighboring vertex of vertex $[N_j]$.
- (e) Output the maximal cliques with “ M_s ” ($s = 1, 2, \dots$) and modify vertex degree of the vertex with label “ $DV = 1$ ” in the adjacency matrix about the given network.
- (f) Go to the step (b) if vertex degree of some vertices is not zero; Go to the step (g).
- (g) If vertex degree of each vertex is zero, quit the repetition, and get all maximal cliques.

3.2. Expand the maximal cliques

In this step, we use the following rules to expand the maximal cliques and detect the community structure.

First, find overlapping vertices belonging to more than one maximal clique with label “OV”. Then, find bridge vertices connecting more than two maximal cliques, not belonging to any maximal clique, and some isolated vertices with low vertex degree are found, which are not belonging to any of the maximal cliques.

Second, N_M is less total number of vertices in the two maximal cliques, and N_O is the total number of the common vertices between these two maximal cliques. These two maximal cliques can be merged into a larger sub-graph when $\frac{N_M}{2} \leq N_O$ in un-weighted networks. Otherwise, these two maximal cliques cannot be merged into a larger sub-graph. However, some weighted networks with different characteristics have different edge weight on behalf of special meaning. So in these weighted networks, two maximal cliques can be merged into a larger sub-graph considered the meaning of weight.

Third, isolated vertices with low vertex degree do not belong to any maximal clique. If an isolated vertex connects with only one non-overlapping vertex, this isolated vertex and vertex connected with it can be divided into a same community. Otherwise, isolated nodes will be not divided into any community.

For clarification, consider the small cooperative network with 17 vertices in Fig. 2, three maximal cliques $M_1 = \{1, 2, 3\}$, $M_2 = \{4, 5, 6, 7, 8\}$, $M_3 = \{14, 15, 16, 17\}$ are extracted. Then vertex 11 is as a bridge vertex, which connects M_2 and M_3 . Vertices 9, 10, 12 and 13 are isolated vertices. According to the above rules, three communities $\{1, 2, 3, 13\}$, $\{4, 5, 6, 7, 8\}$, $\{14, 15, 16, 17\}$, one bridge vertex 11, and three isolated vertices 9, 10, 12 are detected by this proposed algorithm shown in Fig. 3.

3.3. Modularity

In 2004 Newman and Girvan [10] proposed the modularity Q to measure the quality of a particular division of a network:

$$Q = \sum_i (e_{ii} - a_i^2) = Tre - \|e^2\|, \tag{1}$$

where e_{ii} is the fraction of all edges in the network that link vertices in community i to vertices in community j , in which the communities are newly generated. a_i is the fraction of edges which connect to vertices in community i . This modularity Q measures the fraction of the edges in the network that the number of within-community edges minus the expected value of the same quantity in a network with the same community divisions but random connections between the vertices. If the number of within-community edges is no better than random, we will get $Q = 0$. Values approaching $Q = 1$, which is the maximum, indicate obvious community structure. Despite maximum modularity Q corresponds to the partition that comprises the most inner edges within communities and the least outer edges between communities, and the wide acceptance and applications in some real networks, the modularity Q faces several problems. In Ref. [23], it suffers a resolution limit

Table 1
Summary of numerical experimental results.

Data	Karate	Football	Netscience	Cond-mat-2003
Nodes/links	34/78	115/615	1589/2742	31163/120029
GN algorithm	0.416/0 s	0.572/95 s	0.825/138 s	0.424/1256 s
Louvain algorithm	0.431/0 s	0.594/0 s	0.838/0 s	0.457/175 s
Chen algorithm	0.427/0 s	0.585/0 s	0.841/1 s	0.461/192 s
Our algorithm	0.442/0 s	0.601/0 s	0.842/0 s	0.458/180 s

problem and cannot tackle overlapping communities. So, Shen et al. [24] proposed an extended measure Q_c to quantify the overlapping communities.

The modularity in Eq. (1) can be rewritten as

$$Q = \frac{1}{2m} \sum_{c \in P} \sum_{uv} \delta_{cu} \delta_{cv} \left(A_{uv} - \frac{k_u k_v}{2m} \right), \quad (2)$$

where A_{uv} represents an adjacency matrix with the total number of the edge between u and v , P is a partition of network, $k_u = \sum_v A_{uv}$ is the sum of the edges attached to vertex u , $\delta_{cu} = 1$ denotes vertex u belongs to community c and 0 otherwise, $m = \frac{1}{2} \sum_{uv} A_{uv} \lim_{\chi \rightarrow \infty}$ is the total number of edges.

However, considering a vertex might belong to more than one community, we rewrite the modularity in Eq. (2) as

$$Q_m = \frac{1}{2m} \sum_{c \in P} \sum_{uv} \beta_{uc} \beta_{vc} \left(A_{uv} - \frac{k_u k_v}{2m} \right), \quad (3)$$

where

$$\beta_{uc} = \frac{1}{N_{u \in c}} \frac{n_{uc}}{k_u}, \quad (4)$$

and $N_{u \in c}$ is the number of vertex u belonging to communities, n_{uc} is the number of edges adjacent to vertex u in the community, $k_u = \sum_v A_{uv}$ is the sum of the edges attached to vertex u . Obviously, the definition in Eq. (4) satisfies the following conditions

$$0 \leq \beta_{uc} \leq 1, \quad \forall c \in P, u \in V. \quad (5)$$

Furthermore, if u belongs to only one community c , β_{uc} is equal to 1. That is to say, Eq. (4) is consistent with the definition of the modularity for a non-overlapping community in Ref. [10].

4. Results

The proposed detecting overlapping community algorithm is implemented by Java programming language running on a personal computer with 3.0 GHz processor, 4.0 GB memory and Win7 operating system. In order to further verify the performance of our proposed algorithm, we have applied it on some test-case networks, such as Zachary's karate club network [25], football network, collaboration network, and cond-mat-2003. These networks can be downloaded from <http://www-personal.umich.edu/~mejn/netdata/>. In Table 1 we have compared it with three other algorithms (GN algorithm [10], Louvain algorithm [26], Chen algorithm [27]) in networks of various sizes. For each algorithm, the table displays the modularity Q_m and the computation time.

The well-known "karate club" [25] is widely used as a benchmark for some algorithms in detecting community structures. Zachary observed social interactions between the members of a karate club at an American university over the course of two years and constructed this network, which contains of 34 members of the karate club as vertices and 78 edges is representing friendships among members of the club. A dispute arose between the club's administrator and the club's instructor whether to raise club fees, and as a result, the club eventually split in two smaller clubs, centered round the administrator and the instructor respectively. We evaluate the proposed algorithm by testing on un-weighted Zachary's karate club network. First, 25 maximal cliques are extracted from Zachary's karate club during the test of our algorithm. The number of overlapping vertices belonging to two maximal cliques reduces with the merger of two maximal cliques. Only two overlapping 1 and 32 are detected at last. Then four overlapping communities which are represented by the circular, square, triangular and hexagon are detected with Q_m being 0.4118 in Fig. 4. One bridge vertex 10 and one isolated vertex 12 are detected at the same time. This network is also discussed in many papers. Different four communities are detected in Refs. [22,23]. Four non-overlapping communities are detected by the algorithm in Ref. [28]. Two communities and one common vertex with 10 are detected in Ref. [27] with the modularity Q_o is 0.4214. In Ref. [29] two communities are overlapping, sharing a five vertices {3, 9, 10, 14, 30}. As we see in Table 1, although Louvain algorithm is a fast algorithm, it is not able to detect overlapping communities, whereas our algorithm can find them (see Fig. 4).

In May 2006, M.E.J. Newman compiled a weighted collaboration network of scientists who are publishing on the working of network theory and experiment [14]. There are in total 1589 scientists and 2742 coauthor-ships in this collaboration

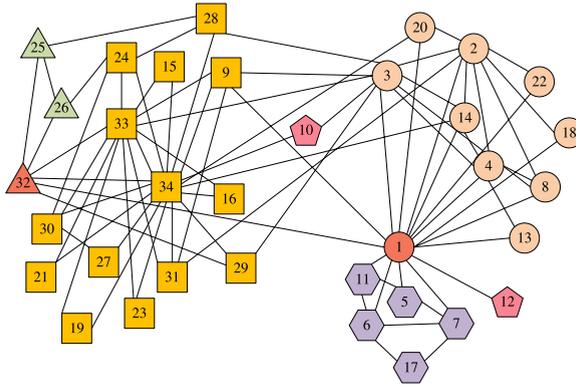


Fig. 4. Four overlapping communities from Zachary's karate club study.

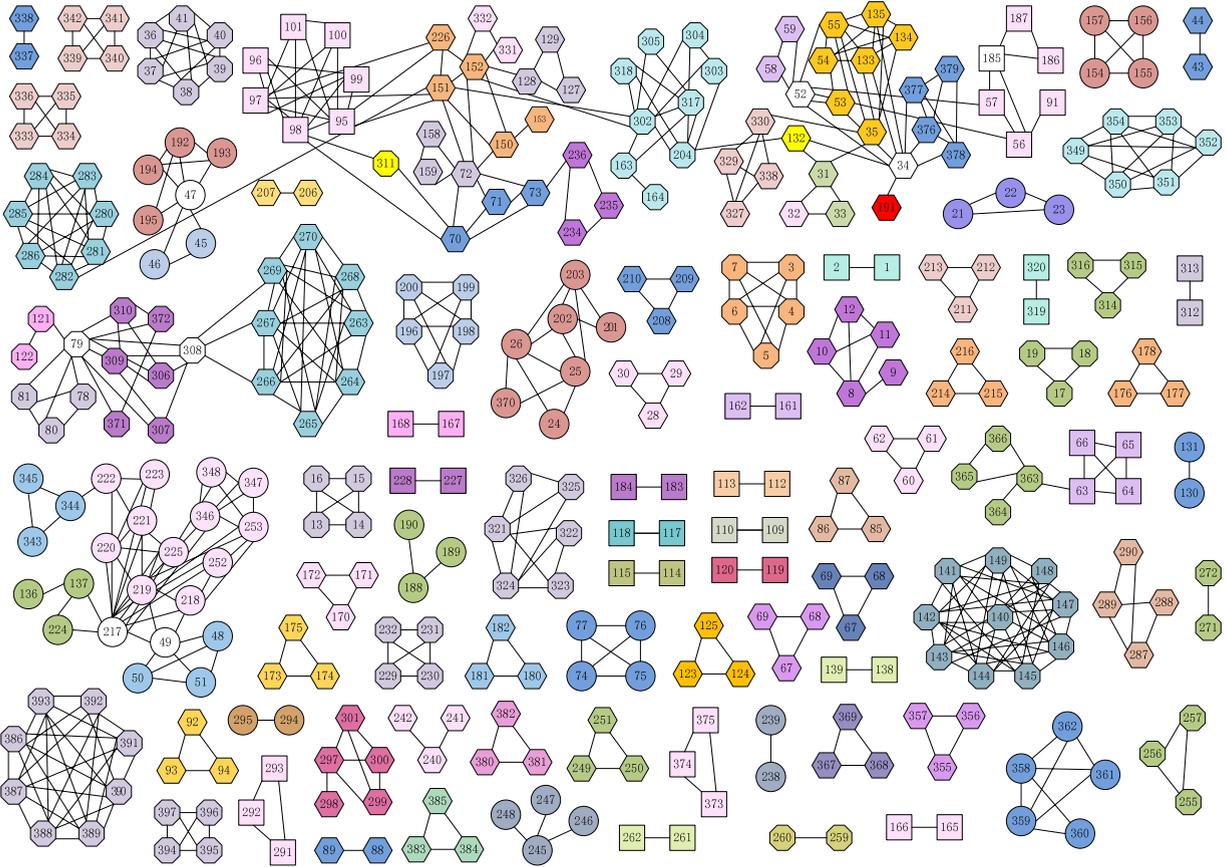


Fig. 5. 98 overlapping communities from Netscience-coauthor network.

network. We just take the largest connected component with 397 scientists as the test benchmark. First 100 maximal cliques are extracted from this collaboration network. Then we use the above rules to expand the maximal cliques and detect the community structure in Section 3.2. 98 overlapping communities are detected from 397 scientists by this proposed algorithm with the modularity Q_m being 0.8411 in Fig. 5. The vertices with white background are overlapping vertex, and many of these overlapping vertices are cross-disciplinary researchers. The vertex 191 is the isolated vertex. The vertices 311 and 132 are the bridge vertices.

We extensively have tested the performance of our algorithm on artificial networks with built-in hierarchical community structure. Each network was constructed with 128 vertices which are split into 4 communities of 32 vertices each [30]. Pairs of vertices falling to the same community are linked with probability p_{in} while pairs falling to different communities are linked with probability p_{out} . The values of p_{in} and p_{out} were chosen to make the expected degree of each vertex equal to 16.

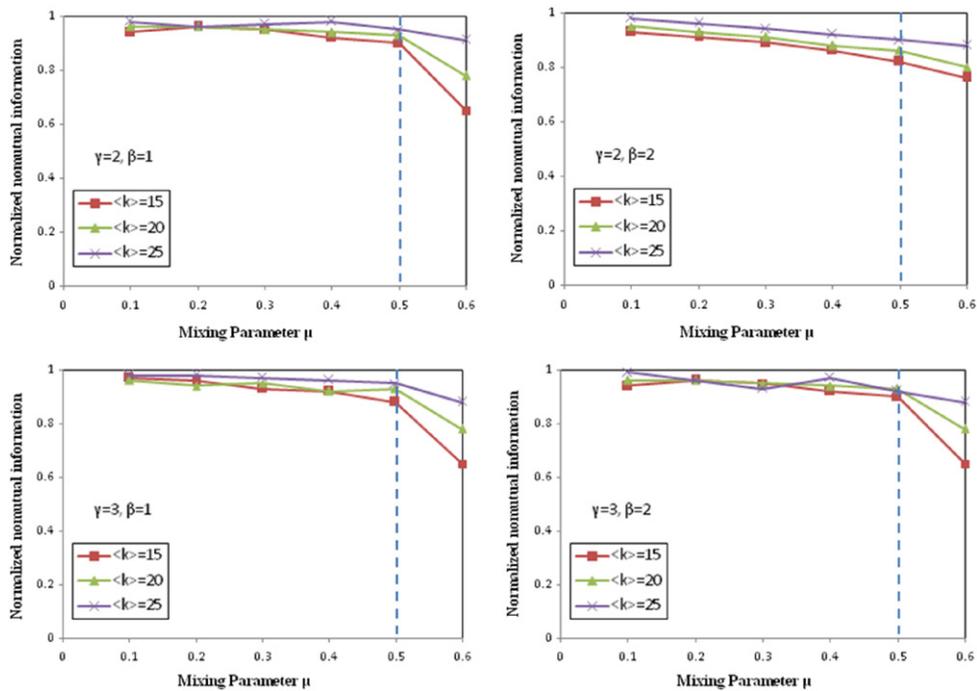


Fig. 6. Test of our algorithm on the LER benchmark. The number of vertices is $N = 1000$. The results clearly depend on all parameters. γ is the exponent for the degree, β is the exponent for the community size distribution, and $\langle k \rangle$ is the average degree.

The accuracy of the method is evaluated by measuring the fraction of correctly identified nodes and the normalized mutual information. The fraction of correctly identified vertices is 0.67 for $p_{\text{out}} = 8$, 0.92 for $p_{\text{out}} = 7$ and 0.98 for $p_{\text{out}} = 6$ in the benchmark proposed in Ref. [30]. In Ref. [10], more than 90% of all vertices are classified correctly from $p_{\text{out}} = 0$ all the way to around $p_{\text{out}} = 6$ by the shortest-path and random-walk versions of the algorithm perform excellently. Only for $p_{\text{out}} \geq 6$ does the classification begin to deteriorate markedly. Our algorithm has been tested on LFR benchmarks proposed in Ref. [29]. We can infer that the proposed algorithm gives satisfactory results from the figures shown in Figs. 6 and 7.

5. Conclusions

In Ref. [7], Palla et al. proposed a clique percolation (CP) algorithm to detect the overlapping community structure. They think that a community can be regarded as a collection constituted by some small completely connected networks in a sense. These completely connected networks can be called clique. k -clique is a completely sub-graph with k vertices. If two k -cliques have $k - 1$ common vertices, these two k -cliques are adjacent. In CP algorithm, cliques are found based on descending order and iterative regression firstly. After all the cliques are found, clique-clique overlap matrix is gotten. Last k -communities can be detected. Unfortunately, CP algorithm fails to detect the primary binary division in this graph. In Ref. [31], T.S. Evans proposed the clique graphs algorithm to detect the overlapping communities. To find communities, the clique graph will be constructed firstly in a weighted graph and then apply a good vertex partitioning algorithm to the clique graph. Various definitions and motivations for these weights are given. The clique graphs denoted by $C^{(3)}(G)$, $D^{(3)}(G)$ and $\bar{D}^{(3)}(G)$ correspond to the adjacency matrices defined in (4), (6) and (7) of Ref. [31] respectively. The highlights and the weakness of the CP method are concluded in Ref. [31].

In summary, a novel algorithm to detect overlapping communities in un-weighted and weighted networks is proposed in this paper. The first pivotal step of this proposed algorithm is extracting all maximal cliques based on deep and bread searching. Then two maximal cliques can be merged into a larger sub-graph by some given rules to expand the maximal cliques and detect the community structure. This proposed algorithm detected in some real networks, which are completely composed of many maximal cliques connected by some overlapping vertices will save more time. In addition, overlapping vertex, bridge vertex and isolated vertex are detected by this proposed algorithm. Two representative real networks are used to test performance of the proposed algorithm. In un-weighted Zachary's karate club network, four overlapping communities are detected with two overlapping vertices, one bridge vertex and one isolated vertex. In weighted collaboration network, 98 overlapping communities are detected from 397 scientists. Some overlapping vertices, bridge vertices and isolated are detected by this algorithm at the same time. Further some large-scale networks, such as football network, collaboration network, and cond-mat-2003 are used to evaluate the overlapping community detection algorithm. We have also tested

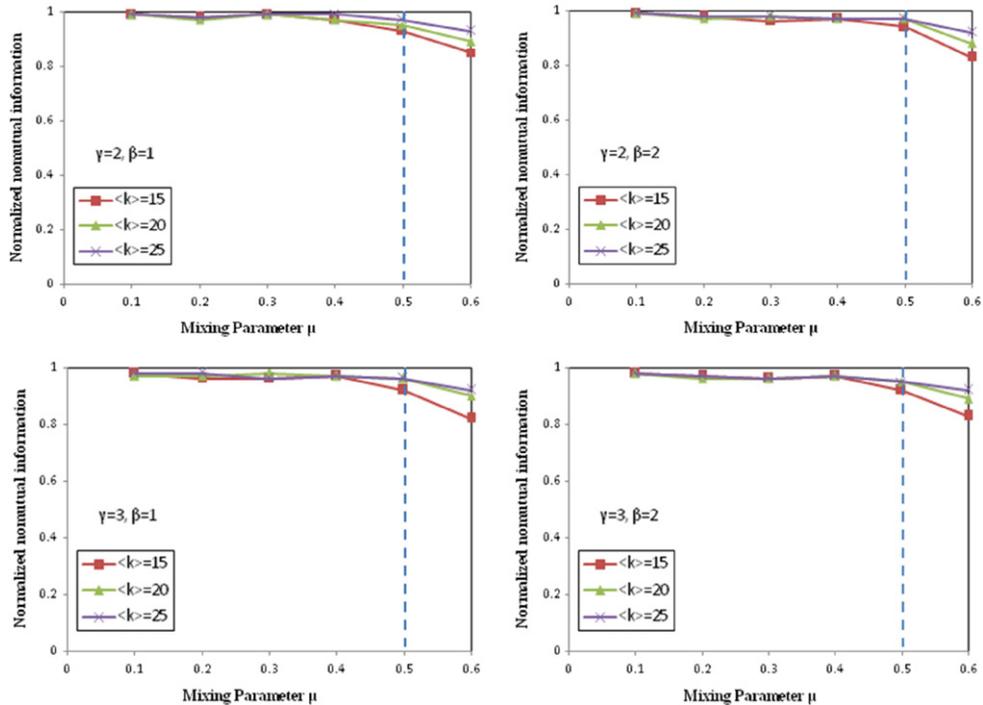


Fig. 7. Test of our algorithm on the LER benchmark with $N = 5000$ vertices. The other parameters are the same as in Fig. 6.

this proposed algorithm on LER benchmark. These experiments demonstrate that the proposed algorithm is rather efficient for detecting the overlapping communities in the complex network.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (Nos. 61370145, 61173183, and 60973152), the Doctoral Program Foundation of Institution of Higher Education of China (No. 20070141014), Program for Liaoning Excellent Talents in University (No. LR2012003), the National Natural Science Foundation of Liaoning province (No. 20082165) and the Fundamental Research Funds for the Central Universities (No. DUT12JB06).

References

- [1] M.E.J. Newman, The structure and function of complex networks, *SIAM Rev.* 45 (2) (2003) 167–256.
- [2] S.H. Strogatz, Exploring complex networks, *Nature* 410 (6825) (2001) 268–276.
- [3] R. Albert, A.L. Barabasi, Statistical mechanics of complex networks, *Rev. Modern Phys.* 74 (1) (2002) 47–97.
- [4] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, *Proc. Natl. Acad. Sci. USA* 99 (12) (2002) 7821–7826.
- [5] S. Fortunato, Community detection in graphs, *Phys. Rep.* 486 (3–5) (2010) 75–174.
- [6] M.E.J. Newman, The structure of scientific collaboration networks, *Proc. Natl. Acad. Sci. USA* 98 (2) (2001) 404–409.
- [7] G. Palla, I. Derényi, I. Farkas, T. Vicsek, Uncovering the overlapping community structure of complex networks in nature and society, *Nature* 435 (9) (2005) 814–818.
- [8] B.W. Kernighan, S. Lin, An efficient heuristic procedure for partitioning graphs, *Bell Syst. Tech. J.* 49 (2) (1970) 291–307.
- [9] M. Fiedler, Algebraic connectivity of graphs, *Czechoslovak Math. J.* 23 (98) (1973) 298–305.
- [10] M.E.J. Newman, M. Girvan, Finding and evaluating community structure in networks, *Phys. Rev. E* 69 (2) (2004) 026113.
- [11] M.E.J. Newman, Fast algorithm for detecting community structure in networks, *Phys. Rev. E* 69 (6) (2004) 066133.
- [12] A. Clauset, M.E.J. Newman, C. Moore, Finding community structure in very large networks, *Phys. Rev. E* 70 (6) (2004) 066111.
- [13] J. Duch, A. Arenas, Community detection in complex networks using extremal optimization, *Phys. Rev. E* 72 (2) (2005) 027104.
- [14] M.E.J. Newman, Finding community structure in networks using the eigenvectors of matrices, *Phys. Rev. E* 74 (3) (2006) 036104.
- [15] X.T. Wang, G.R. Chen, H.T. Lu, A very fast algorithm for detecting community structures in complex networks, *Physica A* 384 (2) (2007) 667–674.
- [16] D.B. Chen, Y. Fu, M.S. Shang, A fast and efficient heuristic algorithm for detecting community structures in complex networks, *Physica A* 388 (13) (2009) 2741–2749.
- [17] X.Y. Lou, A.K.S. Johan, Finding communities in weighted networks through synchronization, *Chaos* 21 (4) (2011) 043116.
- [18] Z.Y. Xia, Z. Bu, Community detection based on a semantic network, *Knowl.-Based Syst.* 26 (2012) 30–39.
- [19] P.G. Sun, L. Gao, S.S. Han, Identification of overlapping and non-overlapping community structure by fuzzy clustering in complex networks, *Inform. Sci.* 181 (6) (2011) 1060–1071.
- [20] H.W. Shen, X.Q. Cheng, K. Cai, M.B. Hu, Detect overlapping and hierarchical community structure in networks, *Physica A* 388 (8) (2009) 1706–1712.
- [21] M.E.J. Newman, Scientific collaboration networks. I. Network construction and fundamental results, *Phys. Rev. E* 64 (1) (2001) 016131.
- [22] R.S. Burt, *Structural Holes: The Social Structure of Competition*, Harvard University Press, Cambridge, USA, 1992.
- [23] S. Fortunato, M. Barthélemy, Resolution limit in community detection, *Proc. Natl. Acad. Sci. USA* 104 (1) (2007) 36–41.
- [24] H.W. Shen, X.Q. Cheng, J.F. Guo, Quantifying and identifying the overlapping community structure in networks, *J. Stat. Mech.* 2009 (2009) 07042.

- [25] W.W. Zachary, An information flow model for conflict and fission in small groups, *J. Anthropol. Res.* 33 (4) (1977) 452–473.
- [26] V.D. Blondel, J.L. Guillaume, R. Lambiotte, Etienne Lefebvre, Fast unfolding of communities in large networks, *J. Stat. Mech.* 2008 (2008) 10008.
- [27] D.B. Chen, M.S. Shang, Z.H. Lv, Y. Fu, Detecting overlapping communities of weighted networks via a local algorithm, *Physica A* 389 (19) (2010) 4177–4187.
- [28] B.S. Rees, K.B. Gallagher, Overlapping community detection using a community optimized graph swarm, *Soc. Netw. Anal. Min.* 2 (4) (2012) 405–417.
- [29] A. Lancichinetti, S. Fortunato, János Kertész, Detecting the overlapping and hierarchical community structure in complex networks, *New J. Phys.* 11 (3) (2009) 033015.
- [30] L. Danon, A. Díaz-Guilera, J. Duch, A. Arenas, Comparing community structure identification, *J. Stat. Mech.* 2005 (2005) 09008.
- [31] T.S. Evans, Clique graphs and overlapping communities, *J. Stat. Mech.* 2010 (2010) 12037.